STUDIES IN DIFFERENTIAL GAMES WITH APPLICATIONS TO OPTIMAL CONTROL UNDER UNCERTAINTY

A thesis submitted
In partial fulfilment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY

by

Rammohan Krishnanand Ragade
to the
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

May 1968

Dedicated to

my Parents

CERTIFICATE

Certified that this work has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

I. G. Sarma

Dr. I.G. Sarma
Assistant Professor
Department of Electrical Engineering
Indian Institute of Technology
Kanpur

May 8, 1968

ACKNOWLEDGEMENTS

I wish to acknowledge explicitly my debt to Professor I.G.Sarma for kindling my imagination with various conceptual problems through his constant questioning and fruitful discussions; to Professors H.K.Kesavan, J.N.Kapur, M.A.Pai, V.Rajaraman for their keen interest, encouragement and useful suggestions. Professor V. Rajaraman, through his lively discussions on problems of adaptive control, and Professor J.N.Kapur through his regular seminars in the Operations Research Group and lectures on Operations Research, have provided some of the motivations of this thesis.

I should also thank my colleagues Mr. U.R.Prasad and Mr. V.V.Mandke with whom I had many useful discussions. Thanks are also due to the departments of Electrical Engineering and Mathematics for providing a stimulating atmosphere for work and the various facilities offered. I should also appreciate the services of Mr. Madhukar Matu for typing this thesis with great interest. Finally I acknowledge the financial assistance offered by the National Institute of Sciences of India, New Delhi during the tenure of this research by the award of a Senior Research Fellowship.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SPECIAL SYMBOLS, NOTATIONS, ETC.

CHARACTER	WORD DESCRIPTION	REMARKS (IF ANY)
(a) Greek Letters:	(Capital letters in parenthesis)	
γ (Γ)	gamma	
δ (Δ)	delta	
η (H)	eta	
θ (Θ)	theta	
ζ (Z)	zeta	
ξ (Ξ)	ksi	
ϕ (Φ)	phi	
ψ (Ψ)	psi	
σ (Σ)	sigma	(used for functions, surface, matrix, summations)
χ	chi	
τ	tau	
μ	mu	
ν	nu	
ρ	rho	
λ	lambda	
(Ω)	omega	
(b) Latin Script:	(Capital letters in parenthesis)	
$\mathscr{b}$ ($\mathscr{B}$)	b (B)	
$\mathscr{c}$ ($\mathscr{C}$)	c (C)	

$(\mathcal{H})$	(H)
$(\mathcal{I})$	(I)
$(\mathcal{K})$	(K)
$m\ (\mathcal{M})$	m (M)
$p\ (\mathcal{P})$	p (P)
$(\mathcal{R})$	r (R)
$(\mathcal{S})$	(S)
$(\mathcal{U})$	(U)
$(\mathcal{V})$	(V)
$(\mathcal{Y})$	(Y)
$(\mathcal{Z})$	(Z)

(c) <u>Circled Letters</u>:

C	C
G	G
m	m
M	M
S	S
U	U
V	V

(d) <u>Superscripts</u>:

$*, o$	denotes optimal points	Other superscripts have usual significance.
	denotes mixed strategies	
i	generally a player	
k	generally discrete time or a player	
m	maximum	

(e) Subscripts:

$*, m$	minimum	Other subscripts have usual significance.

(f) Notations:

$E\{\ \}$, $E(\ )$	Mathematical expectation operator.	
$\langle x, y \rangle$	Bilinear composition	(a) Scalar product if $x \in X$, $y \in X^*$ (b) Composition of a matrix operator and a vector to give another vector.
$\delta_x F$	Frechet differential of F with respect to x	
$\underline{u} = (u^1, \ldots, u^N)$	The collection of strategies of N-players	
$(\underline{u};\ u^i)$	The ith strategy has a different property from the rest of the strategies which have collectively the same property.	
$(\underline{u};\ u^i, u^j)$	The ith and jth strategy differ from the rest.	

(g) Special Remarks:

Unless otherwise stated, we employ vector-matrix notation. Transposes of these quantities are give only where a likely confusion may occur. As far as possible, the special symbols are mutually exclusive from chapter to chapter.

SYNOPSIS

RAMMOHAN KRISHNANAND RAGADE, Ph.D.
Department of Electrical Engineering
Indian Institute of Technology, Kanpur
May 1968
STUDIES IN DIFFERENTIAL GAMES WITH APPLICATIONS TO OPTIMAL CONTROL UNDER UNCERTAINTY

A basic framework for answering many conceptual problems arising in differential games and related areas is developed. In particular, the notion of a generalized differential game, termed the 'Positional Game' is introduced. A 'Positional Game' consists of a system whose 'position' is observed according to different laws by the various players whose actions influence the system. These actions are required to be determined in terms of the observations and inferences only. By considering the different effects of uncertainty in the system description and the nature of information allowed to the players, in the form of memory, observation constraints, unequal information, incomplete and imperfect information, different game problems have been identified. We consider the deterministic game problems in chapters 3 and 4, games with uncertainty in chapters 5 to 7.

In deterministic games, the formulation of linear differential games with constraints is treated in function spaces. The formulation is also examined for games with partial

information. A related problem for linear games is that of playability of strategies and the considerations of controllability, observability and sufficient coordinates. These are also treated. The generalization of two-person differential games to N-person differential games is of recent origin. A heuristic derivation of the necessary conditions for the class of differential games given by Berkovitz is given in the non-cooperative case. This is applied to three problems, viz., the national economy model of two nations, the system design problem with two criteria and the study of 'silent' and 'noisy' 'races'.

In the games under uncertainty, positional games with different information patterns to the players are investigated. In the presence of uncertainty, decisions have to be taken by players under either certainty, risk or ignorance. These require subjective and objective factors to be considered and the implication in on-line and off-line games is examined. The trainer-learner game and the 'dangerous' game are examples of positional games with uncertainty (perhaps ignorance too). The solution and formulation of linear stochastic games with complete and random partial information reveals that a saddlepoint cannot, in general, be satisfied in the latter problem in pure strategies. Further, since each player uses subjective criteria, the appropriate conditions are those obtained in the corresponding N-person non-cooperative theory. The conceptual problems raised in chapter 5 lead to the notion of a two-person Markov Positional Game in chapter 6 and a

general N-person Markov Positional Game in chapter 7. A decomposition of the umpire's state into observed, inferred, remembered and ignored states is established. The properties of mixed, behaviour, equalizer, etc. strategies are examined for the Markov Positional Game. The theory parallels the dual control theory (as developed by Aoki). In the N-person game, the conversion of an incomplete structural information game to one with incomplete position information is shown. The notion of playability is generalized to incorporate certainty, risk and subjective factors. It is shown that a Nash equilibrium point exists if a completely mutually playable tuple exists with strategies which are all stationary and equalizers.

We conclude finally by reiterating the large number of conceptual problems raised above as also the vast potentialities for future work in this area.

I INTRODUCTION

The principal objective of this thesis is to provide a basic framework within which we can answer many conceptual problems that arise in differential games and related problems. In turn, a two-person, zero-sum differential game has been looked upon as a generalisation of the optimal control problem to a situation where direct antagonism exists between two intelligent 'controllers'. Our motivation to introduce the notion of a 'positional game' in chapter 2 is manyfold. Through this notion we show the basic structure in a differential game, whether deterministic, stochastic or with distributed parameters. Secondly, we link it up with the existing literature on extensive games, stochastic games, and recursive games. We thus reduce a gap in this direction. Thirdly, we have made provision for considering games involving decision-making problems, when players are faced with imperfect and incomplete information. We can also bring under the purview of the same framework, games with dynamical constraints that involve several payoffs to the players; games that involve sub-,super-,multi-criteria. The motivation for making these facilities available in a generalised differential game is derived from existing problems in control theory, systems theory and control and systems engineering practice. We are also concerned with problems of adaptive-optimal control problems and optimal control problems under uncertainty with an 'on-line' identification procedure. We feel that the framework provided should enable us to consider these as problems of games and decision theory.

1.2 BRIEF REVIEW OF RELEVANT GAMES

In a dynamical situation, certainly, the most important game is the differential game due to Isaacs|1|, Berkovits|2| and Pontryagin|3|. A two-person zero-sum deterministic differential game consists of the constraints

$$\dot{x} = G\,(x,u,v,t) \tag{1.1}$$

$$K_1\,(x,u,t) \leq 0 \tag{1.2}$$

$$K_2\,(x,v,t) \geq 0 \tag{1.3}$$

$$\phi(x(T)\,,T) = 0 \tag{1.4}$$

and the payoff by player II to player I is given by

$$I(u,v) = \int_{t_0}^{T} f(x,u,v,t)dt + g(x(T),T) \tag{1.5}$$

where x is the 'state', u is the control of player I, v is the control of player II, and ϕ is the terminal constraint. The optimal strategies are those which satisfy the saddlepoint condition

$$I(u^o,v) \geq I(u^o,v^o) \geq I(u,v^o) \tag{1.6}$$

In |3| the optimal strategies of player I are determined after the optimal strategies of player II have been determined. The framework of these two games is in itself not rich enough to provide a general model for considering the various decision problems in which we are interested. Given the above differential game and applying the principle of degradation of uncertainty|4 -5| it is immediately apparent that the above model becomes inadequate to handle several other interesting problems.

The notion of the positional game that is introduced in chapter 2 is an attempt to extract the essential structure in all these problems, and different dynamical games can be defined on the basis of this concept. The hierarchical type of game in |3| has a different information structure. Hence, a study of information structures in differential games would enable us to identify the effect of the principle of degradation of uncertainty in different games. Von Neumann |6| first considered in detail the problem of information and most games discussed in literature are of perfect and complete information. We point out that the notion of a positional game introduced here has no direct connection with the positional game of Milnor|7| and Mycielski|8|.

With the degradation of uncertainty, the differential games have to be considered in a stochastic environment. Kuhn|9|, Dalkey|10| and Davis|11| have considered the information aspect in games which are played sequentially, termed the extensive games. The notions are built upon information sets associated with the game tree. Kuhn showed the profitability of considering games in extensive form to give a richer supply of useful results. The stochastic differential game|12| (and the positional game) turns out to be the counterpart of stochastic|13| and recursive|14| games. These were games with matrix payoffs, played repeatedly over time.

We can consider many control systems problems as decision problems. As the differential game is a two-sided version of the deterministic optimal control problem, the Markov game|15| is the two-sided version of the Markovian decision problem|16|. In the

literature on dynamical games, the two-person zero-sum game concept has received the greatest attention. There are other types of games which have been investigated although not to the same extent. Petrosyan|17|, Karvovskiy and Kuznetsov|18| have considered N-person differential games. Their treatment for games under uncertainty requires further investigation. Let us next review some of the problems of control under uncertainty and the various approaches possible to solve the problem.

1.3 PROBLEMS OF CONTROL UNDER UNCERTAINTY

In the literature on control systems, the effort has always been to provide schemes by which unknown and unreckoned factors have a markedly subdued effect on the performance of the system. The various criteria by which such schemes were evaluated have been, in general, in terms of wide platitudes. The advent of optimal control theory and modern control techniques has changed the outlook and viewpoints in problem-solving. There have, however, been reported curious blends of the modern and classical techniques which, though it has evolved practical schemes, have yet to find a place in modern control theory|19|. One such concept, that of adaptation, borrowed from biological sciences has kindled the imagination of not only the writers of science fiction stories, but also various workers in control literature. The resulting 'pot pouri' has led to an indiscriminate use of terms such as 'adaptation', 'learning', 'self-organisation', etc.|20|. The central theme of all these cases revolves around a basic uncertainty in the problem description and the desire to

have the best system operation according to a prespecified criterion despite the uncertainty.

On examination it is seen that all adaptive processes are characterised by the existence of a state space and an information space with transformations on them satisfying the semigroup property|21| and these transformations are such as to optimally remove uncertainty while optimally controlling the process(system). Thus an adaptive-optimal control process is essentially a bidecision-making process: one decision is made to minimise the cost of control and the other to minimise the cost of estimation or identification. Both the decisions are made on-line. The very presence of uncertainty requires a new approach to the problem. We feel that a solution to this problem lies in the general framework of games and decisions.

At this point we wish to examine in brief a different approach through the theories of learning and pattern recognition. Learning in control systems has three requirements|22|:

(i) An experiment organized into a sequence of identical trials which must be performed.

(ii) Each trial must produce some performance or output by the system.

(iii) The performance must be measurable and a relationship better than defined for the performance scores.

This in turn requires a repetitive nature of the desired situation. Some generalisation of the existing theory of recursive and stochastic games in this direction could provide an alternative framework. Its applicability to control problems is limited to

regulator type of problems with periodic repetition of command signals. The application of pattern recognition techniques requires classification of the uncertainty into discrete subsets and the use of an on-line identification technique to determine these subsets.

Since in general most control problems can be shown to be decision-making problems, whether with certainty, risk or uncertainty, the corresponding generalisation to games must incorporate these features. A one-sided decision-making problem with uncertainty is often viewed as a game against nature.

With extraneous constraints in the form of memory, time, control inputs, computability, etc., there arise further decision-making problems. The problem of on-line optimisation under uncertainty is one such decision-making problem. For adaptive-optimal control problems, Sworder|23| has given a rigorous formulation for the rational synthesis of adaptive-optimal controllers blending Wald's philosophy with that of dual control theory|24|. The adaptive-optimal control problem is shown to exist in problems which have a dominant set of control policies, none of which are strictly preferable over each other. In such situations the role of supercriteria is obvious. Some of these concepts have been rigorously defined by Witsenhausen|25|. No such problems have been considered for control situations with more than one decision-maker.

We next consider a brief outline of the various chapters.

1.4 OUTLINE OF VARIOUS CHAPTERS

The notion of a positional game has been introduced to consider N-person decision-making in general, in the absence of complete information in a dynamical environment. A positional game for N-players is defined as the collection of sets

$$\left(\overline{\{U^i\} , X , W, \{Y^i\}, S, \{H_i\}, \Theta}, \{I^i\} \right)$$

where the overbarred sets represent the system and $\{I^i\}$ the payoff functions, specified in terms of $x \in X$, the positions, and $\{u^i\} \in \{U^i\}$, the control actions of each player. W is the set of all parameters, S the system function, $\{H_i\}$ is the observation function of each player, $i=1,\ldots,N$. W is correctly the cartesian product of subparameter sets one of which is W_d the set of parameters characterising absolute uncertainty. A deterministic differential game has the structure

$$\left(\overline{\{U^i\} , X, \{Y^i\} , S, \Theta}, \{I^i\} \right)$$

along with a prespecified information structure. The 'position' and 'state' coincide here. In almost all other cases the player has to enlarge his 'position vector' to a 'hyperposition' or state. In chapter 2 we consider this aspect and many others taking into account the memory, information and constraints on observations and inputs. In chapter 3 we consider the formulation aspects of differential games in function spaces and show that this normal form of study becomes inadequate to consider the extensive game properties of memory, linkage of information sets etc. For games with partial information the problem of

controllability, observability and sufficient coordinates is examined. With two or more persons in the game with multiple payoffs, the concept of a N-person differential game is introduced in chapter 4. A set of necessary conditions for a tuplet of strategies to be at a Nash equilibrium point are derived. In chapter 5, games under uncertainty are considered with provision to account for problems with human factors. The solution of linear stochastic differential games is next given. Chapter 6 and chapter 7 point a way to consider games with incomplete information which have barely been investigated in literature. The notion of a Markov Positional Game is a generalization of the dual control concept of Fel'dbaum|24| to the two-person complete structural information case. In chapter 7 more properties of N-person Markov Positional Games are examined. The notion of a playable pair of strategies is generalized.

It is thus felt that the positional game approach to differential games is a partial answer towards a comprehensive theory of games with incomplete information. This meets an idea due to Bellman|26| where he points out that it is the engineer's job to achieve partial control with partial information and partial specifications.

II PROBLEMS OF POSITIONAL GAMES

The concept of a positional game is introduced to generalise the differential game model. As pointed out in |2|, not all games termed differential games in literature are indeed so. Some of these cannot be termed games in the strict sense. In fact it is possible to discover many other games which have similar structure but have other information and decision structures. Our attempt in this chapter is, therefore, directed towards considering the various properties of these allied games through a basic structure. We shall use many ideas from existing control theory and decision theory. The various notions will be introduced in the form of definitions followed by explanatory remarks.

2.2 NOTATIONS, DEFINITIONS AND ASSUMPTIONS

This section follows partly |28| in giving precise definitions of some fundamental concepts.

<u>Definition 2.1</u> : The system S is a collection of objects $(U,V,W,X,Y,Z,S,H_1,H_2,\Theta)$ in which U,V,W,X,Θ are nonempty sets and

$$\begin{aligned} S&: U \times W \times V \times \Theta \times X \to X \\ H_1&: X \times W \times \Theta \to Y \qquad (2.1) \\ H_2&: X \times W \times \Theta \to Z \end{aligned}$$

U is termed the <u>control set</u> of player I

V is termed the <u>control set</u> of player II

$u \in U$ is termed a <u>control action</u> for player I

$v \in V$ is termed a <u>control action</u> for player II

$w \in W$ is termed a <u>parameter</u>

$t \in \Theta$ is termed a <u>time</u>. Θ is totally ordered

$x \in X$ is the <u>position</u> of the game

$y \in Y$ is the <u>observation</u> of player I

$z \in Z$ is the <u>observation</u> of player II

S is the <u>system function</u>, H_1 is the <u>observation function</u> of player I, H_2 is the <u>observation function</u> of player II.

<u>Definition 2.2</u> : The system is <u>one-sided</u> if the control set of exclusively one player is a singleton (equivalence class). The system is <u>N-sided</u>, or with <u>N-persons</u> or <u>with N-decision makers</u>, if the collections U, V, can be extended to $\{U^i\}$, $i = 1,\ldots,N$, Y, Z, to $\{Y^i\}$, $i = 1,\ldots,N$, and H_1, H_2 to $\{H_i\}$ $i = 1,\ldots,N$.

Henceforth we shall consider the N-sided system.

<u>Assumption 2.1</u> : The control action of player i will be a function of the parameter $w \in W$, a time interval $T \subset \Theta$, and the observation set $y^i \subset Y^i$.

<u>Assumption 2.2</u> : We use the notation W to denote product of <u>subparameter</u> sets $W = W_1 \times W_2 \times \ldots \times W_q$.

[This allows consideration of the distributed parameter system or stochastic systems].

<u>Assumption 2.3</u> : In general the observation functions are not the same. This leads us to define

<u>Definition 2.3</u> : The game is said to be of <u>equal observation</u> if and only if $H_i = H_j$, $y^i \subset Y^i$, $y^j \subset Y^i$ for all $i, j = 1,\ldots,N$, $i \neq j$. In all other cases the game is said to be of unequal observation.

From assumption 2.1 we are led to define

<u>Definition 2.4</u> : The set Γ_u^i (S) of all control actions of player i , is the set of all functions γ_u^i such that $\gamma_u^i : Y \times W \times \Theta \to U^i$ which have the property that for each fixed $w \in W$, $t \in \Theta$, the equations

$$u^i = \gamma_u^i \, (\, (S \, (\underline{u}, t, w, x), \, w, t \,), \, w, t \,) \tag{2.2}$$

has a unique solution for each fixed u^j, $j \neq i$, and x, in u^i for all $i, j = 1, \ldots, N$.

<u>Definition 2.5</u> : The <u>payoff function</u> or the <u>objective function</u> of the ith player is the map $I^i : X \times U^1 \times U^2 \times \ldots \times U^N \times \Theta \times W \to \mathbb{R}$. An element $p^i \in I^i$ is termed the <u>payoff</u>.

<u>Definition 2.6</u> : The N-person game is then the tuplet (S , γ_u^1 , γ_u^2 , ..., γ_u^N, I^1, I^2, ..., I^N).

[<u>Remark</u>: If each decision maker has more than one criterion, we can certainly replace the single decision maker with many criteria by many decision makers with single criterion.]

The N-person game is depicted in Fig. 2.1.

<u>Definition 2.7</u> : Let W be a parameter set with a probability distribution over the elements of W. We term such games <u>stochastic games.</u>

[<u>Remarks and observations</u>: Stochastic games were first considered by Shapley |13|; stochastic differential games |12| and stochastic duels |27| are other stochastic games. In our stochastic game the stochastic element enters via (i) the system function, (ii) the observation function of each player, (iii) the randomization resorted to by the players, (iv) in the uncertainty of the payoff and finally (v) in the nature of

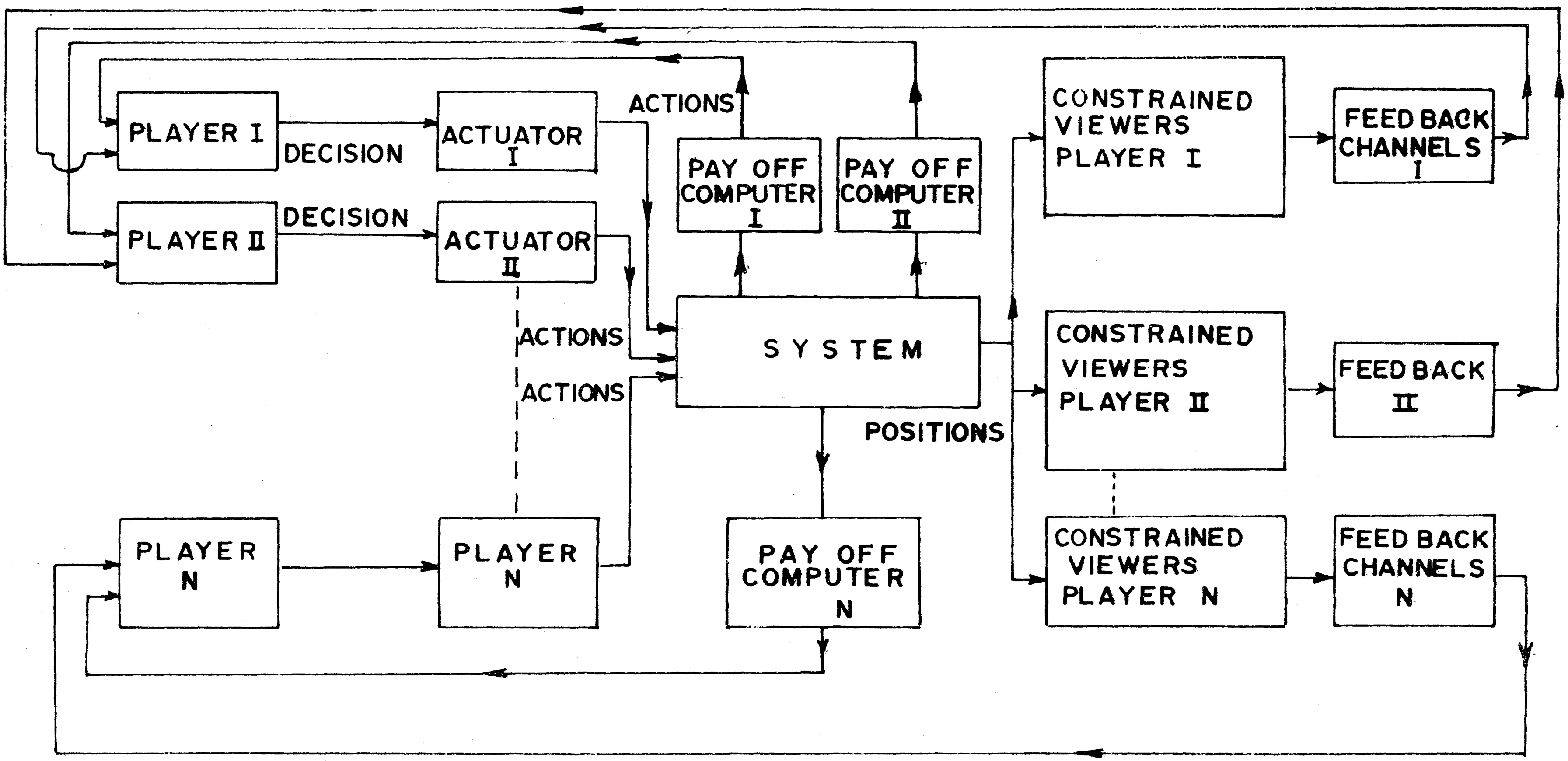


FIGURE 2·1. AN N-PERSON POSITIONAL GAME.

a player's knowledge of the resources and capabilities of the other players.]

<u>Definition 2.8</u> : A stochastic game is said to have <u>fixed statistics</u> if the distribution of the statistics of the parameters is fixed and time independent.

If the distribution is known and time varying the stochastic game has instead a fixed specified <u>stochastic process</u>.

[Both the above games can be converted into stochastic differential games.]

Where player i has only an apriori knowledge of the distribution, this becomes an <u>adaptive stochastic game</u>.

<u>Definition 2.9</u> : Let the parameter W be identified with some spatial coordinates. Then we say that this is a <u>distributed positional game</u> or the <u>field game</u>.

<u>Definition 2.10</u> : If the boundaries of the field game are fixed it is said to have <u>fixed domain</u>.

If the boundaries are time varying, the field game is said to have <u>variable domain</u>.

[<u>Remarks</u>: 1. The close similarity between a field game and a stochastic game is seen in definitions 2.9 and 2.10.

2. The position as yet is not identified with terms such as state. This is done in the next section. The position is there some physical entity whose knowledge is essential but which is only partially observed. Note that the position vector is always finite.

3. We do not consider field games in this study.]

2.3 STATE, POSITION AND MARKOV PROPERTY

Much of classical physics hinges on the concept of causality. The behaviour over time of a dynamical system is looked upon as a continuous sequence of transformations |21|. The 'state' of the system at time t_o is a system function whose knowledge at time t_o together with the sequence of transformations (semigroup of transformations |28-29|) suffices to determine the state at any future time. The knowledge of the 'state' behaviour prior to t_o is contained in the 'state' at t_o. For many deterministic systems, with finite number of phase coordinates, it is natural to identify the phase (or a linear transformation of it) with the state. However, if to each phase there correspond many future trajectories dependent on an undeterminable set of parameters, such an identificat is not possible. In such a case a precise definition of the notion of state is called for. In quantum mechanics where a similar problem arises, the dilemma is resolved through the introduction of a new system function satisfying the properties of state |30|.

It is true that for many physical systems alternative approaches to define the notion of state and to give a state description of the system are possible |31|. For systems whose input-output behaviour is completely given, well-defined methods are possible. In the positional game much depends on the information structure. This framework should be general enough to permit consideration of certain types of ill-defined problems as arise in chapter 5.

Supposing we require that the 'state' for a player refer to the minimum amount of packaged information necessary for future evaluations of strategies, payoff etc. In this framework we make

provision for situations where players have a decision-making problem under uncertainty. Thus, a player's a priori 'beliefs' are also candidates for 'state'. The terms 'state' and 'position' are distinguished as follows. Each player has the same position while each player can have a different state space, and the constraints and specifications themselves are given only in terms of the positions.

Before we discuss the properties of state and position further, we need a few properties of the set Θ.

<u>Lemma 2.1</u> : Let $t_1 \in \Theta$. Consider $\Theta' = \Theta - \{t_1\}$. Then $\Theta' = \Theta'_- \cup \Theta'_+$, $\Theta'_- \cap \Theta'_+ = \phi$.

<u>Lemma 2.2</u> : Given $\vartheta \subset \Theta$ there exists a unique $t_o \in \bar{\vartheta}$ and $t_f \in \bar{\vartheta}$ such that for all $t \in \bar{\vartheta}$ $t_o \leq t$ and $t \leq t_f$.

<u>Proof</u> : The proofs of these lemmas follow from the well-ordered properties of Θ.

<u>Definition 2.11</u> : The set Θ'_- is the set of all <u>past times</u> of t_1 or Θ'_- constitutes the <u>past</u> of t_1.

<u>Definition 2.12</u> : The set Θ'_+ is the set of all <u>future times</u> of t_1 or Θ'_+ constitutes the <u>future</u> of t_1.

<u>Definition 2.13</u> : We shall term 't_o' the <u>initial time</u> and 't_f' the <u>final time</u>.

[These concepts are similar to those introduced for infinite games by Kuhn |9|, Dalkey |10| among others. The element t_o is the distinguished vertex of the tree |9|].

Assumption 2.1 now allows us to define 'control functions'.

<u>Definition 2.14</u> : The <u>control function</u> of player i is the map $\mu^i_T : T \times W \times Y \rightarrow U^i$ where $T \subset \Theta$. The subscripts T indicates the

<u>Definition 2.15</u> : The <u>range of control functions</u> of player i is the set $\mathcal{U}_T^i$ of the maps, $\mu_T^i(w)$ for all $w \in W$. Let $\Omega_{u^i} = \underset{T \in \Theta}{x} \mathcal{U}_T^i$. Then $U^i \subset \Omega_{u^i}$.

<u>Definition 2.16</u> : The <u>past control functions</u> at $t_i \in \Theta$ is denoted by μ_-^i while the <u>future control function</u> is denoted by μ_+^i. The <u>past system function</u> at $t_i \in \Theta$ is denoted by S_-, while the <u>future system function</u> is denoted by S_+.

<u>Definition 2.17</u> : A '<u>state</u>' is a function of the game which has the property of separating past from future, with the future depending only on the present state.

[Remark: This definition does not constitute a working definition with well set properties for the 'state'. The past and future times as well as the control functions are depicted in Fig. 2.2. In classical deterministic dynamical systems if $S(t_o)$ stands for the initial state and $u[t_o,t]$ represents a forcing function over the interval $[t_o,t]$ then the entire future is contained in $S(t)$ where $S(t)$ is the state obtained from $S(t_o)$ through the transformations $\phi(t,t_o)$. This transformation has the group property

$$\phi(t,t_o) = \phi(t,t_1)\,\phi(t_1,t_o) \quad \text{for all} \quad t_1 \in R \tag{2.3}$$

For causal systems we restrict this to the semigroup

$$\phi(t,t_o) = \phi(t,t_1)\,\phi(t_1,t_o) \quad \text{for all} \quad t_1 \in [0,\infty) \tag{2.4}$$

In the context of our game the state of a player has to behave such that the transformations from state to state constitute a semigroup. When some components of the state vector represent the player's knowledge based on observations and information, the only possible way the semigroup property

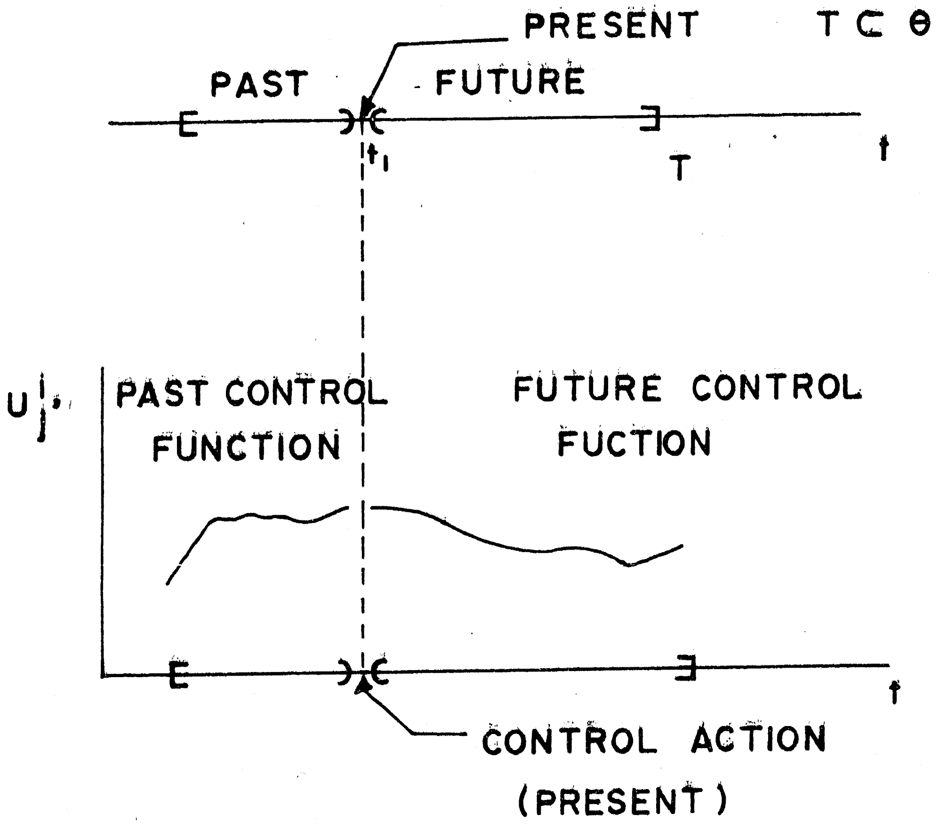


FIGURE 2.2 PAST, PRESENT, AND FUTURE-CONTROL FUNCTIONS.

could be true is by requiring the state to be Markovian. Thus the Markovian property is linked with the concept of state. Clearly the position is always a subset of the state.]

While the position is a specified entity it is upto the player to completely define the state vector. The process of adding variables to the position vector to form the state vector is similar to the process of inflation in Dalkey |10|. Supposing we have inflated the position vector to form the state function space Ψ. Then we test Ψ according to the following definition |28|

<u>Definition 2.10</u> : Given the state function space Ψ^i for player i, the control sets $\{U^i\}$, the parameter set W the system position set X, the observation sets $\{Y^i\}$ and a subset $T \subset \Theta$. We have for each $u^i \in U^i$ a continuous mapping $\phi_s : T \times \Psi \times T \to \Psi$ with the property that

(i) $$\phi_s(t_2;\ \phi_s(t_1,\ \Psi_o^i(w), t_o),\ t_1) = \phi_s(t_2;\ \Psi_o^i(w), t_o) \qquad (2.8)$$

for all $t_o < t_1 \leq t_2$ in T and all $\Psi_o(w) \in \Psi$.

(ii) $\phi_s(t_1;\ \Psi_o(w),\ t_o) \to \Psi_o(w)$ as $t_1 \to t_o$

for all $t_o \in T$ and all $\Psi_o(w) \in \Psi$.

(iii) The functions f, H_1, H_2 are continuous with respect to x.

We shall always consider the minimal function which satisfies this as the state.

2.4 MEMORY, INFORMATION AND CONSTRAINTS

This section is devoted to the study of the 'rules' of the game, that is, various constraining maps on the sets of the

system. Less rigorously each player is equipped with limited resources in the form of limited data processing equipment, finite memory, constraints on the observations and inputs, etc. The rules can differ from player to player. Some of the concepts are defined precisely while a heuristic discussion seems appropriate for the others. In a realistic situation, it is unfair to expect a player to have completed all his computations instantaneously if only due to limited computational speed and the amount of data he can handle. This in turn reflects on his ability to recall all his past observations and control actions at any given instant. The definitions below are for an arbitrary player i. The tuplet under consideration will be (U^i, Y^i, W, Θ).

<u>Definition 2.19</u> : Let $m : U^i \times \Theta \times Y^i \times W \to R^{a^i} \times \Theta$ where $a^i = \text{rank } U^i + \text{rank } Y^i$. Then m shall be called the <u>action history of player i</u>.

<u>Definition 2.20</u> : Let $H \subset m$, $H: U^i \times T \times Y^i \times W \to R^{a^i} \times T$, $T \subset \Theta$, then H is an <u>action history</u> for time T.

<u>Definition 2.21</u> : Let $\underline{H} \subset m$. $\underline{H}$ is said to be a <u>memory</u> for player i.

<u>Definition 2.22</u> : Let $\underline{H} \supset H$. Then the memory is <u>unlimited for time T</u>. If this is true for all T the <u>memory is unlimited</u>.

<u>Definition 2.23</u> : If T is a connected set then the action history is called a <u>recall $\mathcal{R}$</u> $_T$.

<u>Lemma 2.3</u> : A recall $\mathcal{R}_T$ has unlimited memory if the memory and recall have the same set T.

<u>Lemma 2.4</u> : A memory $\underline{H}$ is connected if there exists a connected

set T such that $\mathcal{R}_T \subseteq M$ and $M \subseteq \mathcal{R}_T$.

Proof: Proof of Lemma 2.3 is obvious. To prove 2.4 we assume the contrary. Then since m is also a memory over θ, $\mathcal{R}_\theta$ and m are identical. $\mathcal{R}_\theta = m$. Let M be connected. Then the action history is connected which implies the set T over which the action history is defined is connected. Q.E.D.

Definition 2.24 : If in every recall $\mathcal{R}_T$ the corresponding T includes t_o, the initial time, then the recall is perfect.

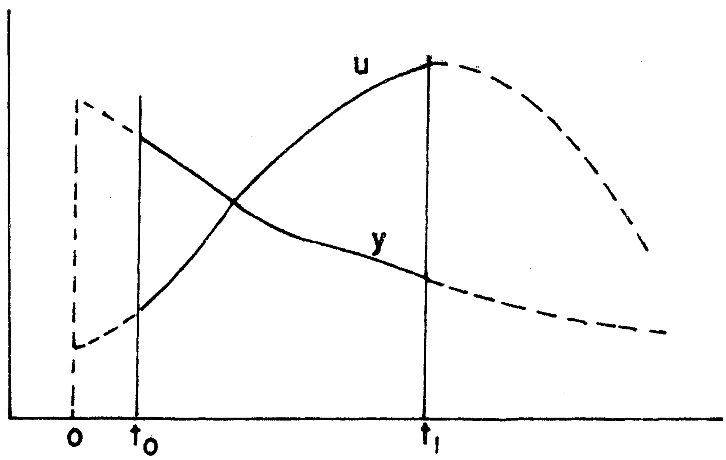
Definition 2.25 : If a recall $\mathcal{R}_T \subset M$, then the memory M is sufficient for T. If every subset $T \subset \theta$ that includes t_o the recall $\mathcal{R}_T \subset m$, then the player is said to have perfect recall. The corresponding memory is termed perfect memory.

Lemma 2.5 : If for any time T that includes t_o the player i has perfect memory, then he has perfect recall.

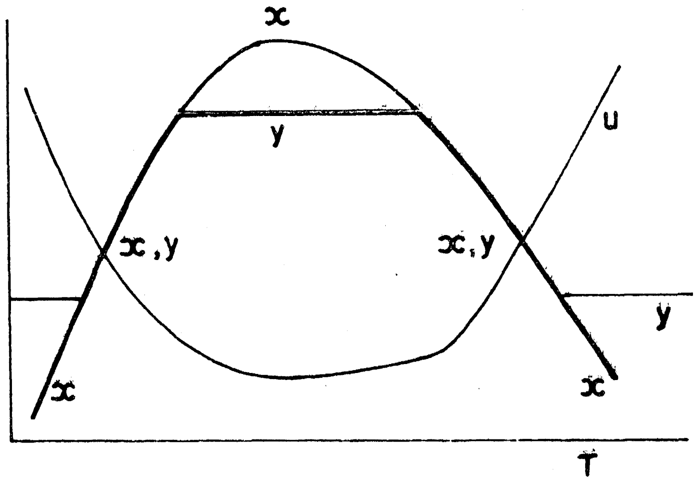
[In Figs. 2.3 - 2.5 we depict the notions of memory, recall, perfect recall and sufficient memory].

We now turn our attention to the problem of information which appears in a variety of ways. In chapter 12 of Isaacs |1| there is a discussion of various possibilities. Many more are possible in an N-person game. The rules of the game prescribe what types of information are permissible for the players. We distinguish two types of information |32, 35|.

Definition 2.26 : The game is said to have incomplete information if at least to one player the rules of the game allow some indeterminacy in the payoffs, system functions and of constraints, if any.

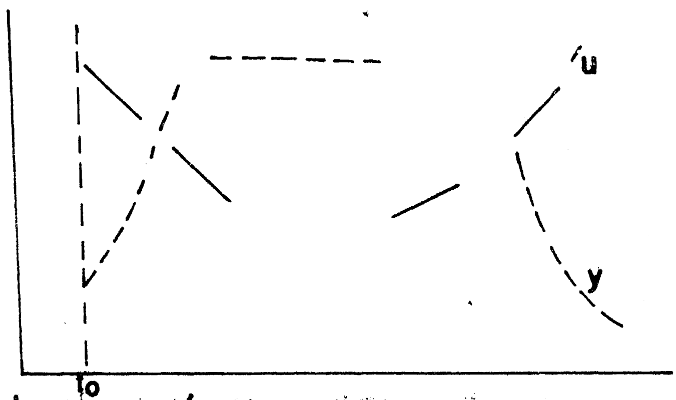


ACTION HISTORY FOR TIME (t_0, t_1)

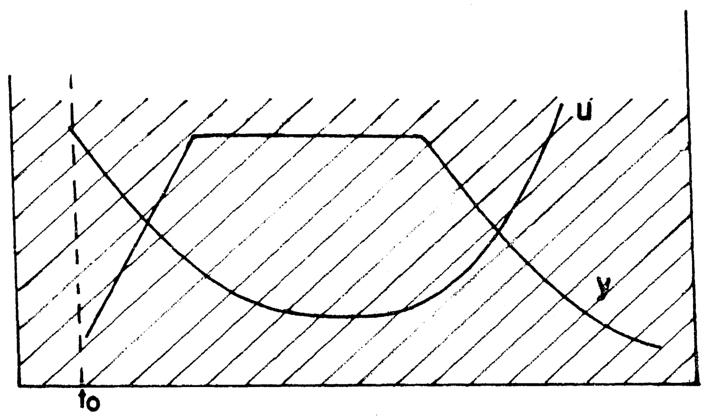


CONSTRAINED OBSERVATIONS

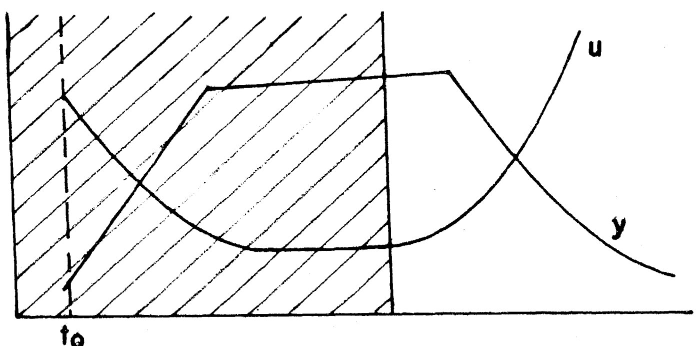
FIGURE 2.3 CONSTRAINED OBSERVATIONS AND ACTION HISTORY.



A 'MEMORY' OF OBSERVATIONS AND INPUTS



SUFFICIENT MEMORY



INSUFFICIENT MEMORY

FIGURE 2.4 MEMORY SUFFICIENT AND INSUFFICIENT.

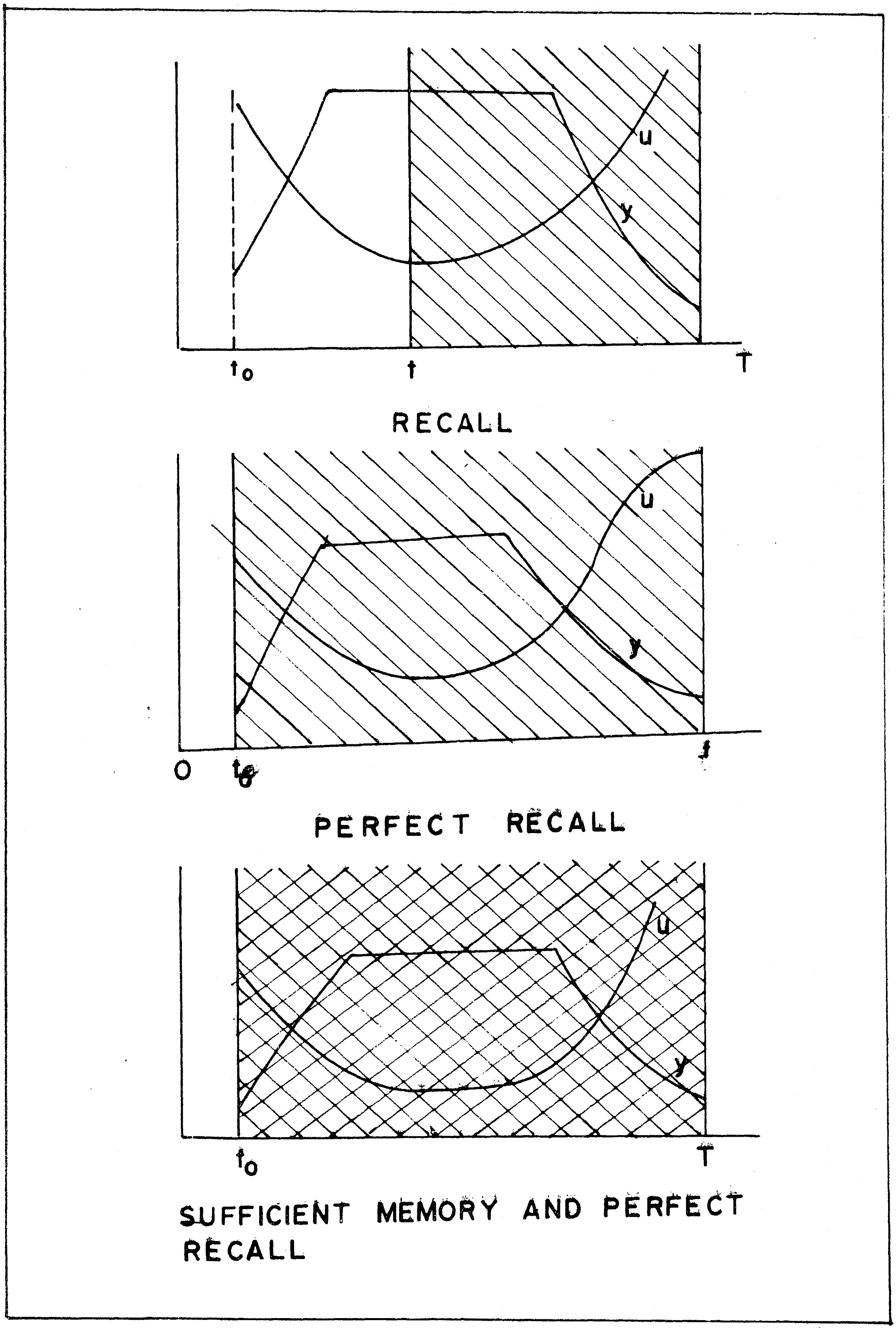


FIGURE 2·5 RECALLS.

<u>Definition 2.27</u> : The game is said to be of <u>imperfect information</u> when player i is uncertain about his own previous moves and observations, his knowledgeof other players' moves and observations.

[Observe that imperfect information arises due to memory constraints while incomplete information arises due to partial knowledge of the structure of the game].

These two main classes have various other subclasses.

<u>Definition 2.28</u> : The players have <u>complete</u> information if there is no uncertainty in the form of the nature of parameters and constrained observations, and when all players have equal observations equal to the position of the game.

<u>Definition 2.29</u> : The players are said to have <u>partial information</u> if some position variables are suppressed in the observations.

<u>Definition 2.30</u> : The game is said to be of <u>incomplete structural information</u> to player i, if the exact nature of the system or observation functions or payoff functions or all i are ill-defined to him.

<u>Definition 2.31</u> : The player i is said to have <u>null information</u> if he has no observations. Such a game to him is a <u>search game</u>

<u>Definition 2.32</u> : All observations constitute <u>granted</u> information of a player.

<u>Definition 2.33</u> : When the game to a player is of <u>perfect recall</u>, then he has <u>perfect information.</u>

<u>Definition 2.34</u> : Any information obtained through data processing of the action history constitutes <u>inferred information.</u>

__Definition 2.35__ : Any information exchanged between player i and player j is termed the __communicated information__ between player i and player j.

__Theorem 2.1__ : Let every player be granted perfect information, complete and equal information. Let every player resort to no randomization. __Then the game is such that the position and state coincide.__

__Proof:__ This follows, since under perfect information no uncertainty can exist. Under complete and equal information the observation function H_i = identity operator, and, therefore, $y^i = x$. Further, since the position itself is a suitable candidate for the state, there is no need for any player to do additional information processing and randomize his control functions. Thus $u^i = u^i(x,t)$ Q.E.D.

The above theorem is satisfied by all deterministic differential (difference) games [1]. One source of imperfect information is in having discrete observations on a continuous process.

We now consider other types of constraints.

__Definition 2.36__ : An input constraint is a mapping

$$K^i : U^i \times W \times Y^i \times \Theta \to \mathcal{R}^{m^i} \times \Theta .$$

__Definition 2.37__ : An observation constraint is a mapping

$$H^i : X \times W \times Y^i \times \Theta \to \mathcal{R}^{l^i} \times \Theta .$$

[__Remarks__ : When considering a stochastic game, the constraints can be specified in terms of expected quantities or with a specified distribution].

__Definition 2.38__ : A player is said to perform __on-line tasks__ if his control action is dependent on the previous actions and

observations generated while the game proceeds.

Definition 2.29 : A player is said to perform off-line tasks if his control action is a prescription for all future times determined at the start.

Games with pre-play communication among players are off-line games.

[Note: In the two-person differential game considered by Isaacs |1| and Berkovitz |2|, the players have perfect, complete and equal information and, therefore, the players can determine their strategies either on-line or off-line. In the games considered by Kirillova |34|, Pontryagin |3| the strategy of player II has to be prespecified to player I who then chooses his strategy. This game can be imbedded in the following game. Let N players in a N-person game be ranked. Player 1 chooses his strategy at any given instant first, player 2 then chooses his strategy knowing what player 1 has chosen. Then player 3 chooses his strategy knowing what 1 and 2 have chosen and so on. The Nth player chooses last of all. Inherent in a deterministic game approach is the requirement of instantaneous communicated information from the players lower up in the hierarchy to the higher ups. A more realistic formulation should take account of inherent variable time lags in the communicated information and imperfect information. A further step would be to relax the complete information for persons lower in the rank. This class of games we term hierarchical games and has obvious connection with hierarchical control systems].

2.5 CONCLUSIONS

We have described certain notions in formulating positional games and laid them in the form of definitions. In the succeeding chapters we consider only a partial application of all these notions to differential games in function spaces, N-person differential games, stochastic games and Markov positional games.

III FUNCTION SPACE FORMULATION AND LINEAR POSITIONAL GAMES

In studying positional games, so as to obtain the Value (the optimal payoff) and the optimal strategies, it is immediately apparent that each player has two possible approaches. One would be to study the game a priori and expect all other players also to do the same. He then determines an optimal strategy (it is a program here) under the assumption that all players choose optimal strategies for all future times. Such an approach constitutes a normalisation of the game. In such a game the player is only able to set the optimal program and hope that the game would proceed accordingly. In much of the earlier development of game theory, the attempt had been to convert all extensive games to normal form games. Kuhn|9| showed the way to an alternative study. In an almost similar manner there have been two trends in the development of differential game theory. Many optimal control problems |34 - 35|, pursuit evasion and differential game problems |36 - 39| have been studied using functional analysis. It is emphasised here that this is a normal form study. Not all extensive games can be studied in normal forms. Games with incomplete and imperfect information can be reduced to normal form games only under special conditions.

In this chapter a study of normal form games (mostly deterministic differential games) is attempted. We shall try and extend the study to a few positional games with incomplete information. In the dynamic continuous case the use of functional

analysis is convenient, since a number of strategies can be represented by a single element from an appropriate function space. Such use of functional analysis has been made by several authors in the context of pursuit-evasion games, attrition games |39|, optimal control problems |40|. This section is based on the report |41|. In the last section we consider the concepts of controllability, observability and sufficient coordinates in games with incomplete information.

3.2 FORMULATION OF LINEAR DIFFERENTIAL GAMES

We shall formulate herein two typical games:

(a) A Tracking Game: Two nations A and B are at war. A has sent a missile on a predetermined course carrying a warhead to destroy a certain target in B. In the simplified model considered here it is assumed that the missile receives guidance through a radio link and hence is able to interfere with the guidance scheme through jamming. In this manner B tries to prevent the motion of the missile along its preset course. A wishes to achieve its mission with minimum cost of error and control, while B wishes to maximize the same. [This problem is rather oversimplified. In practice the guidance of missiles is not done thus, nor can jamming be achieved in a purely analog manner. This examples only serves to illustrate the formulation.]

Let $x(t)$ denote the n-vector of positions of the missile.

$u(t)$ denote the r-vector of control actions of A

$v(t)$ denote the s-vector of control actions of B.

The governing system equations are assumed to be

$$\dot{x}(t) = A(t)\,x(t) + B(t)u(t) + C(t)v(t) \tag{3.1}$$

where $A(t)$, $B(t)$, $C(t)$ are $n \times n$, $n \times r$, $n \times s$ matrices. Alternatively (3.1) can be written in the integral form

$$x(t) = \phi(t,t_o)x(t_o) + \int_{t_o}^{t} \phi(t,\tau)B(\tau)u(\tau)d\tau + \int_{t_o}^{t} \phi(t,\tau)C(\tau)v(\tau)d\tau \tag{3.2}$$

where $\phi(t,t_o)$ is the fundamental matrix

$$\frac{d\phi(t,t_o)}{dt} = A(t)\,\phi(t,t_o) \quad , \quad \phi(t_o,t_o) = I \tag{3.3}$$

and $x(t_o)$ is the initial position. Let $x_d(t)$ be the desired trajectory and the error defined by

$$e(t) = x(t) - x_d(t) \tag{3.4}$$

Let $\mathcal{H}_x^n, \mathcal{H}_u^r, \mathcal{H}_v^s$, denote Hilbert spaces whose elements are square integrable over $[t_o, T]$. For each $u \in \mathcal{H}_u^r$, $v \in \mathcal{H}_v^s$ we define

$$Lu = \int_{t_o}^{t} \phi(t,\tau)\,B(\tau)\,u(\tau)d\tau$$

$$= \int_{t_o}^{T} \phi(t,\tau)\,B(\tau)\,u(\tau)d\tau \tag{3.5}$$

with $\phi(t,\tau) = 0$ for $\tau > t$, owing to physical realisability requirement (semigroup property |29|). L is a bounded continuous operator, $L:\mathcal{H}_u^r \to \mathcal{H}_x^n$. One can similarly define

$$Fv = \int_{t_o}^{T} \phi(t,\tau)\,C(\tau)\,v(\tau)d\tau \tag{3.6}$$

with $\phi(t,\tau) = 0$ for $\tau > t$. F is a bounded continuous operator, $F:\mathcal{H}_v^s \to \mathcal{H}_x^n$. The equations (3.2) and (3.4) can be now written

$$\left.\begin{aligned} x &= x^o + Lu + Fv \\ e &= Lu + Fv - g \\ g &= x^o - x_d \end{aligned}\right\} \tag{3.7}$$

$$x^0 = \Phi(t,t_0)\, x(t_0) \tag{3.8}$$

The loss to player A is given by

$$I(u,v) = \int_{t_0}^{T} (\langle e(t),\, Q(t)\, e(t)\rangle + \langle u(t),\, R(t)u(t)\rangle - \langle v(t),\, S(t)v(t)\rangle)dt \tag{3.9}$$

which can be written

$$= \langle e,\, Qe\rangle + \langle u, Ru\rangle - \langle v, Sv\rangle \tag{3.10}$$

where we define as the scalar product in

$$\langle p,q\rangle = \int_{t_0}^{T} p(t)\, q(t)dt \tag{3.11}$$

(b) <u>A Pursuit Evasion Game</u>: Two fleet belonging to two warring nations are sailing in international waters. Nation A which has some unusually modern weapons wants them to stay far away from the surveillance of the crafts of B, while nation B has just the opposite interests. The fleet have the linearised dynamics

$$\begin{aligned} \dot{x}_a(t) &= A_a(t)\, x_a(t) + B(t)\, u(t) \\ \dot{x}_b(t) &= A_b(t)\, x_b(t) + C(t)\, v(t) \end{aligned} \tag{3.12}$$

where x_a and x_b are the phases of the two crafts, u the r-control vector, v the s-control vector of B. The objective of A will always be to keep the cumulative squared deviations maximum. Thus the payoff is written

$$I(u,v) = \int_0^T \langle (x_b(t) - x_a(t)),\, (x_b(t) - x_a(t))\rangle dt \tag{3.13}$$

and

$$P(u^0,v^0) = \max_{u\in U}\ \min_{v\in V}\ P(u,v) = \min_{v\in V}\ \max_{u\in U}\ P(u,v) \tag{3.14}$$

As before we can write

$$x_a = L_a u + x_a^o$$

$$x_b = L_b v + x_b^o \tag{3.15}$$

which is the operator representation of (3.12), $L_a : \mathcal{H}_u^r \to \mathcal{H}_x^n$, $L_b : \mathcal{H}_v^s \to \mathcal{H}_x^n$, $x_a^o, x_b^o \in \mathcal{H}_x^n$, and

$$\begin{aligned} I(u,v) &= \langle (x_a - x_b), (x_a - x_b) \rangle \\ &= \langle (L_a u - L_b v), (L_a u - L_b v) \rangle + \langle (x_a^o - x_b^o), (x_a^o - x_b^o) \rangle \\ &\quad + 2 \langle (L_a u - L_b v), (x_a^o - x_b^o) \rangle \end{aligned} \tag{3.16}$$

It is thus seen that these two games are constituted by a set of linear constraints and a nonlinear functional as payoff. In general linear differential games have the following structure. Constraints:

$$x = x^o + Lu + Fv \tag{3.17}$$

$x, x^o \in \mathcal{B}_x^n$, $u \in \mathcal{B}_u^r$, $v \in \mathcal{B}_v^s$, where $\mathcal{B}^n, \mathcal{B}_u^r, \mathcal{B}_v^s$ are Banach spaces, $L : \mathcal{B}_u^r \to \mathcal{B}_x^n$, $F : \mathcal{B}_v^s \to \mathcal{B}_x^n$ and the payoff function

$$P(u,v) = \|u\|_1 - \|v\|_2 + \|x\|_3 \tag{3.18}$$

where the norms are differently specified. The determination of strategies in the above two problems is straightforward. We, however, relate these games to the general positional game problem.

3.3 GAMES IN FUNCTION SPACES

Consider the product Banach spaces $\mathcal{X}, \mathcal{U}, \mathcal{V}$ of rank n, r, s respectively where $x \in \mathcal{X}$ is the position, $u \in \mathcal{U}$ is the control

strategy of I, $v \in \mathcal{V}$ is the control strategy of II. Let X_f and X_o be additional spaces such that

$$\partial_{t_o} \mathcal{X} = X_o \qquad \partial_{t_f} \mathcal{X} = X_f \tag{3.19}$$

where ∂_t is the boundary operator at $t \in \tau$. The system function can then be written

$$x = S(x,u,v) \tag{3.20}$$

and the payoff function we can write as

$$I(u,v) = f(x,u,v) \tag{3.21}$$

where S and f are continuous functions in x,u,v. The specification of the game would be complete with the specification of the rules of the game. We make the following assumptions

(i) only an action-history for finite time is considered.

(ii) an unlimited memory with perfect recall to the players

(iii) no observation constraints.

The input constraints are, however, specified as follows:

For player I:

$$K_1(u,y) \geq 0 \tag{3.22}$$

For player II:

$$K_2\ (v,z) \leq 0 \tag{3.23}$$

where $K_1 \in \mathcal{X}_1$, $K_2 \in \mathcal{X}_2$. $\mathcal{X}_1$, $\mathcal{X}_2$ are suitably defined Banach spaces of rank m and 1 respectively and the relation $x \geq y$ $x \in \mathcal{B}$, $y \in \mathcal{B}$ implies $x-y \in P$, where P is a closed convex cone.

Fig. 3.1 shows the mappings on function spaces to form the Positional Game.

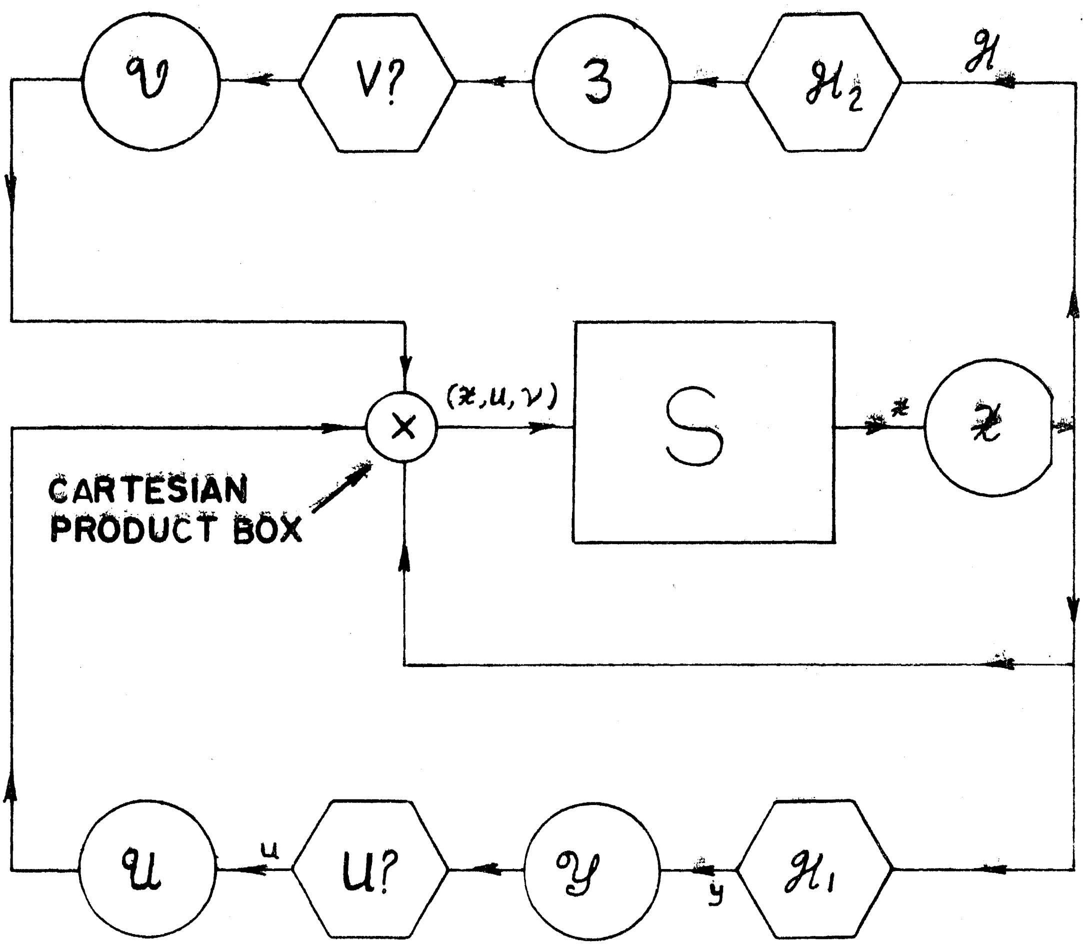


FIGURE 3·1 MAPPINGS OF FUNCTION SPACES TO FORM THE POSITIONAL GAME.

(a) <u>The Complete Information Case</u>: In this case

$$x = y = z \tag{3.24}$$

Let $A \subseteq \mathcal{H} \times \mathcal{U} \times \mathcal{V}$ be such that $a \in A$, $a = (x, u, v)$, a satisfies constraints (3.19), (3.22 - 3.23). Let $\mathcal{U} \cap A = U$, $\mathcal{V} \cap A = V$. Further at the terminal time consider an n-dimensional manifold M in X_f. If for each $u \in U$, every $v \in V$ is such that $x(t_f) \in X_f$, then the pair (u, v) is said to constitute a playable pair. Since no randomisation over u, v are considered here, the playable pair is a set of pure strategies. If in the class of all playable pure strategies (U, V) the saddlepoint inequality holds:

$$I(u^o, v) \leq I(u^o, v^o) \leq I(u, v^o) \tag{3.25}$$

then the game has a saddlepoint pure strategies. We further assume that the saddle value written Value is always unique for a given $x(t_o) \in X_o$ if v_1^o, v_2^o are two candidates such that

$$I(u^o, v_1^o) = I(u^o, v_2^o) \tag{3.26}$$

and
if u_1^o, u_2^o are optimal candidates such that

$$I(u_1^o, v^o) = I(u_2^o, v^o) \tag{3.27}$$

We consider the related program when one player is made inactive, say $v = v^o$.

$$\text{Min} \left\{ I(u, v^o) : u \in U,\ (x, u, v^o) \text{ is feasible} \right\}.$$

Let us form the Lagrangian

$$F = I(u, v^o) + \langle \lambda, S(x, u, v^o) - x \rangle + \langle \mu, K_I \rangle + \langle \nu, K_{II} \rangle \tag{3.28}$$

where $\lambda \in \mathcal{X}^*$, $\mu \in \mathcal{X}_1^*$, $\nu \in \mathcal{X}_2^*$, * denotes the conjugate space.

<u>Assumption 3.1</u>: The functions G, K_1 , K_2 are at least once differentiable in the Frechet sense.

To determine the appropriate conditions we make use of the following theorem in linear spaces |48|.

<u>Theorem 3.1</u> : Let $\mathcal{X}$ be a linear space, $\mathcal{Y}$, $\mathcal{K}$, be linear topological spaces, P_Y, P_K convex cones in $\mathcal{Y}$ and $\mathcal{K}$ with nonempty interiors. X a fixed convex subset of $\mathcal{X}$, f a concave function on X to $\mathcal{Y}$. g a concave function X to $\mathcal{K}$. Let $x^0 \in X$ be such that $g(x^0) \geq 0$, $x^0 \in$ int P_K. If x^0 maximizes $f(x)$ subject to $g(x) \geq 0$, $x \in X$, then there exists linear continuous functionals $\eta \in \mathcal{Y}^*$, $\mu \in \mathcal{K}^*$, $\eta \geq 0$, $\mu \geq 0$, such that for the Lagrangian expression

$$\phi(x,\mu) = \langle \eta , f(x) \rangle + \langle \mu , g(x) \rangle \tag{3.29}$$

the saddlepoint inequality

$$\phi(x,\mu^0) \leq \phi(x^0,\mu^0) \leq \phi(x^0,\mu) \tag{3.30}$$

holds for all $x \in X$ and all $\mu \geq 0$. If one further assumes Frechet differentiability of the functionals then

$$\begin{aligned} &\langle \delta_x \phi (x,\mu^0) , h \rangle \geq 0 \quad h \in X \\ &\langle \delta_\mu \phi (x^0,\mu) , K \rangle \leq 0 \quad K \in P_K \end{aligned} \tag{3.31}$$

This theorem applied to the one person game considered above leads to the following theorem.

<u>Theorem 3.2</u> : Let $u^0 \in U$, $v^0 \in V$, (u^0,v^0) is a playable pair and u^0

a local minimum point, v^0 a local maximum point. Then there exist linear continuous functionals $\lambda^0 \in \mathcal{X}^*$, $\lambda^0 \neq 0$, $\mu^0 \in \mathcal{X}_1^*$ $\mu^0 \leq 0$, and convex neighbourhoods $N(\lambda^0)$, $N(\mu^0)$, such that F also has a local minimum and F satisfies a saddlepoint inequality in $F(x,u,\lambda,\mu)$; with the differentiability of F assumed in the Frechet sense we have

$$\begin{array}{ll} \langle \delta_x F, x_1 \rangle \geq 0 & \langle \delta_u F, u_1 \rangle \geq 0 \\ \langle \delta_\lambda F, \lambda_1 \rangle \leq 0 & \langle \delta_\mu F, \mu_1 \rangle \leq 0 \end{array} \tag{3.32}$$

for all $x_1 \in N(x^0) \subseteq \mathcal{X} \cap A$, $u_1 \in N(u^0) \subseteq \mathcal{U}$, $\lambda_1 \in N(\lambda^0)$, $\mu_1 \in N(\mu^0)$.
Let us look at the implications of (3.32). From (3.19), (3.22 - 3.24) and (3.32) we have

$$\langle (\delta_x f + \langle \lambda, \delta_x (S-I) \rangle + \langle \mu, \delta_x K_1 \rangle + \langle \nu, \delta_x K_2 \rangle), x_1 \rangle \geq 0 \tag{3.33}$$

and

$$\langle \delta_u f + \langle \ , \delta_u S \rangle + \langle \ , \delta_u K_1 \rangle , u_1 \rangle \geq 0 \tag{3.34}$$

In order that λ, μ be unique it is necessary that $(\delta_u K_1)^{-1}$ exists. Similarly $(\delta_x S-I)^{-1}$ exists. These two conditions have come to be known as constraint qualifications.

We can now consider the game when both the players are active. In this case in order that the saddlepoint condition (3.25) remains valid it should be immaterial which player chooses his strategy first. Then we apply twice theorem 3.2, each time keeping one of the players' strategy fixed at his optimal strategy. A set of linear continuous functionals (multipliers) λ^0, μ^0, ν^0, are obtained at the optimal point, which must be unique. We are thus led to

Theorem 3.3 : Let $u^o \in U$, $v^o \in V$, (u^o,v^o) a playable pair and (u^o,v^o) satisfy the saddlepoint inequality (3.25). Then there exist linear continuous functionals $\lambda^o \in \mathcal{X}^*, \lambda^o \neq 0, \mu^o \in \mathcal{K}_1^*, \mu^o \leq 0$, $\nu^o \in \mathcal{K}_2^*, \nu^o \geq 0$ and convex neighbourhoods of $N(\lambda^o)$, $N(\mu^o)$ and $N(\nu^o)$ such that F also satisfies a saddlepoint inequality in (x,u,v,λ,μ,ν). If further the differentiability in the Frechet sense is assumed for the functions S, K_1, K_2, I then

$$< \delta_x F, x_1 > \geq 0 \qquad x_1 \in N(x^o) \tag{3.35}$$

$$< \delta_u F, u_1 > \geq 0 \qquad u_1 \in N(u^o) \tag{3.36}$$

$$< \delta_v F, v_1 > \leq 0 \qquad v \in N(v^o) \tag{3.37}$$

$$< \delta_\lambda F, \lambda_1 > \leq 0 \qquad \lambda_1 \in N(\lambda^o) \tag{3.38}$$

$$< \delta_\mu F, \mu_1 > \leq 0 \qquad \mu_1 \in N(\mu^o) \tag{3.39}$$

$$< \delta_\nu F, \nu_1 > \geq 0 \qquad \nu_1 \in N(\nu^o) \tag{3.40}$$

Corresponding sufficiency conditions can be derived under the restricted assumption of the function $I(u,v)$ being concave in v and convex in u. Correspondingly we can write

$$F(x^o,\lambda^o,u^o,\mu,v^o,\nu^o) \leq F(x^o,\lambda^o,u^o,\mu^o,v^o,\nu^o) + <\delta_\mu F,\mu-\mu^o> \tag{3.41}$$

$$F(x^o,\lambda^o,u,\mu^o,v^o,\nu^o) \geq F(x^o,\lambda^o,u^o,\mu^o,v^o,\nu^o) + <\delta_u F,u-u^o> \tag{3.42}$$

$$F(x^o,\lambda^o,u^o,\mu^o,v,\nu^o) \leq F(x^o,\lambda^o,u^o,\mu^o,v^o,\nu^o) + <\delta_v F,v-v^o> \tag{3.43}$$

$$F(x^o,\lambda^o,u^o,\mu^o,v^o,\nu) \geq F(x^o,\lambda^o,u^o,\mu^o,v^o,\nu^o) + <\delta_\nu F,\nu-\nu^o> \tag{3.44}$$

As an example consider the differential game problem. Here we have the constraints given by

$$x(t_o) + \int_{t_o}^{t} G(x,u,v,t)dt - x(t) = 0 \tag{3.45}$$

$$K_1(x,u,t) \geq 0 , \quad K_2(x,v,t) \leq 0 \tag{3.46}$$

Payoff function:

$$I(u,v) = \int_{t_o}^{T} f(x,u,v,t)\, dt \tag{3.47}$$

The equation (3.45) is the integral form of the differential constraint:

$$\dot{x} = G(x,u,v,t) \tag{3.48}$$

The Lagrangian can now be formed as

$$\begin{aligned} F(x,\lambda,u,\mu,v,\nu) = & \int_{t_o}^{T} f(x,u,v,t)dt + \int_{t_o}^{T} \mu(t)K_1(x,u,t)dt \\ & + \int_{t_o}^{T} \nu(t)K_2(x,v,t)dt + \int_{t_o}^{T} x(t_o)\lambda(t)dt \\ & + \int_{t_o}^{T}\lambda(t) \int_{t_o}^{t} G(x,u,v,s)dsdt - \int_{t_o}^{T}\lambda(t)\, x(t) \end{aligned} \tag{3.49}$$

From theorem 3.3 we write

$$\delta_x F = 0 = \left\{\int_{t_o}^{T} \frac{\partial}{\partial x} (f+\mu K_1+\nu K_2)dt + \int_{t_o}^{T} \lambda(t) \frac{\partial}{\partial x} G(x,u,v,t)dt - \lambda(t)\right\} \delta x$$

$$\Rightarrow \qquad \lambda^{\cdot}(t) = - \frac{\partial}{\partial x} (f+\lambda G+\mu K_1+\nu K_2) \tag{3.50}$$

Similarly

$$\begin{aligned} &\delta_u F = 0 \Rightarrow \mu K_1 = 0 \quad \mu \leq 0 \\ &\frac{\partial}{\partial u} (f + \lambda G + \mu K_1) = 0 \end{aligned} \tag{3.51}$$

$$\delta_v F = 0 \Rightarrow \nu K_2 = 0 \quad \nu \geq 0$$
$$\frac{\partial}{\partial v}(f + \lambda G + \nu K_2) = 0 \tag{3.82}$$

The transversality condition is similarly obtained by considering

$$\delta_t F + \delta_x F = 0 \tag{3.83}$$

at the terminal time which leads to

$$(f + \lambda G)\, dt - \lambda dx = 0 \tag{3.84}$$

for all variations dt, dx along the terminal surface.

(b) <u>The Partial Information Case</u>: We next consider the partial information case. Here the position is known only through observations obtained by the mapping $\mathcal{H}_1 : \mathcal{X} \to \mathcal{Y}$, for player I where $\mathcal{Y}$ is the Banach space of observation of player I, $\mathcal{H}_2 : \mathcal{X} \to \mathcal{Z}$, for player II where $\mathcal{Z}$ is the Banach space of observations of player II. In this case the definition of a strategy is as in definition 2.4. We now seek to find a suitable operator which does the map $u: \mathcal{Y} \to \mathcal{U}$, and $v: \mathcal{Z} \to \mathcal{V}$. Let us suppose functional operators u and v exist. Further suppose that for each $u \in U$, $v \in V$ is playable and for each $v \in V$, $u \in U$ is playable. Then if the operators u, v are bounded piecewise continuous we have the following modification of theorem 3.3.

<u>Corollary 3.4</u>: Let $u^0 \in U$, $v^0 \in V$, (u^0, v^0) a playable pair and (u^0, v^0) satisfy the saddlepoint inequality. Then there exist continuous linear functionals (multipliers) $\lambda^0 \in \mathcal{X}^*$, $\lambda^0 \neq 0$, $\mu^0 \in \mathcal{K}_1^*$, $\mu^0 \leq 0$, $\nu^0 \in \mathcal{K}_2^*$, $\nu^0 \geq 0$ and convex neighbourhoods

of $N(\lambda^0)$, $N(\mu^0)$, $N(\nu^0)$ such that F also satisfies the saddlepoint inequality in $(x, u, v, \lambda, \mu, \nu, y, z)$ where $y \in Y$ and $z \in Z$ are the observations of player I and player II respectively. If further differentiability in the Frechet sense is assumed for S, H_1, H_2, K_1, K_2, I, then

$$\langle \delta_x F, x_1 \rangle = 0 \qquad \text{for all } x_1 \in N(x^0)$$

$$\begin{aligned} \Rightarrow \delta_x F = 0 = \; & \delta_x f + \delta_u f \circ \delta_y u \circ \delta_x y + \delta_v f \circ \delta_z v \circ \delta_x z \\ & + \langle \lambda, ((\delta_x S - I) + \delta_u S \circ \delta_y u \circ \delta_x y + \delta_v S \circ \delta_z v \circ \\ & + \langle \mu, (\delta_x K_1 + \delta_u K_1 \circ \delta_y u \circ \delta_x y) \rangle \\ & + \langle \nu, (\delta_x K_2 + \delta_v K_2 \circ \delta_z v \circ \delta_x z) \rangle \end{aligned} \tag{3.55}$$

and the rest of the conditions remain the same as in theorem 3.3.

Remarks: It is thus seen that whereas in the open loop (program) problems

$$\delta_x F = 0 \Rightarrow \delta_x f + \langle \lambda, (\delta_x S - I) \rangle + \langle \mu, \delta_x K_1 \rangle + \langle \nu, \delta_x K_2 \rangle = 0 \tag{3.56}$$

in the closed loop problems via observations (control) additional factors have to be accounted for the dependence of u on y and y on x, of v on z and z on x. Synthesis via observations will be possible if it is possible to determine functionals as u(y), v(z). This synthesis procedure can be useful if it takes account of proper mappings such as sufficient coordinates which reduce computational work. We shall solve in the next section the game formulated in section 3.2, and in the following section, examine in more detail this question of sufficient coordinates for

3.4 SOLUTION TO A LINEAR GAME

Let us consider the game :

$$\underset{u \in U \subseteq \mathcal{B}_u^r}{\text{Min}} \quad \underset{v \in V \subseteq \mathcal{B}_v^s}{\text{Max}} \quad (\langle x, Qx \rangle + \langle u, Ru \rangle - \langle v, Sv \rangle) \tag{3.57}$$

subject to

$$x = x^0 + Lu + Pv \tag{3.58}$$

and

$$U = \{ u : \|u\| \leq 1 \} \tag{3.59}$$

$$V = \{ v : \|v\| \leq 1 \} \tag{3.60}$$

where $x \in \mathcal{B}_x^n$ and $\mathcal{B}_x^n, \mathcal{B}_u^r, \mathcal{B}_v^s$ are appropriate Banach spaces. $L : \mathcal{B}_u^r \to \mathcal{B}_x^n$, $P : \mathcal{B}_v^s \to \mathcal{B}_x^n$. Let $L^* : \mathcal{B}_x^n \to \mathcal{B}_u^r$, $P^* : \mathcal{B}_x^n \to \mathcal{B}_v^s$ denote conjugate linear transformations corresponding to the continuous linear transformations L, P. We write the constraints (3.59 - 3.60) in a different manner. Define

$$g_1(u) = 1 - \|u\| \qquad g_1 \in \mathcal{R} \tag{3.61}$$

$$g_2(v) = \|v\| - 1 \qquad g_2 \in \mathcal{R} \tag{3.62}$$

Then the game can be equivalently stated as: find elements $u \in \mathcal{B}_u^r$, $v \in \mathcal{B}_v^s$ such that the constraints (3.58, 3.61, 3.62) are satisfied and the payoff function in (3.57) has a saddlepoint. Let the Lagrangian be

$$\begin{aligned} F(x,\lambda,u,\mu,v,\nu) = {} & \langle x, Qx \rangle + \langle u, Ru \rangle - \langle v, Sv \rangle \\ & + \mu (1 - \|u\|) + \nu (\|v\| - 1) \\ & + \lambda (x^0 + Lu + Pv - x) \end{aligned} \tag{3.63}$$

where $\mu, \nu \in \mathcal{R}$, $\lambda \in \mathcal{B}_x^{n\,*}$. Now apply theorem 3.3. If u^0, v^0

form policies) then

$$\delta_x P = 0 \Rightarrow Qx - \lambda = 0 \tag{3.64}$$

$$\delta_u P = 0 \Rightarrow Ru + L^*\lambda - \mu u^* = 0 \tag{3.65}$$

$$\delta_v P = 0 \Rightarrow -Sv + P^*\lambda + \nu v^* = 0 \tag{3.66}$$

where $u^* \in \mathcal{B}_u^{r\,*}$, $v^* \in \mathcal{B}_v^{s\,*}$ are such that $\langle u^*, u \rangle = \|u\|$, $\langle v^*, v \rangle = \|v\|$ and $\|u^*\| = 1$, $\|v^*\| = 1$ by the Hahn-Banach theorem |37|, |43|.

Also

$$\mu(1 - \|u\|) = 0 \tag{3.67}$$

$$\nu(\|v\| - 1) = 0 \tag{3.68}$$

which imply

$$\begin{aligned} &\mu \leq 0 \text{ for } \|u\| \leq 1 \\ &\nu \geq 0 \text{ for } \|v\| \leq 1 \qquad \text{and} \\ &\mu = 0 \text{ for } \|u\| \geq 1 \\ &\nu = 0 \text{ for } \|v\| \geq 1 \end{aligned} \tag{3.69}$$

Operate (3.65) by R^* from the left and (3.66) by S^* from the left, then

$$(R^*R)u + R^*L^*\lambda - \mu R^*u^* = 0 \tag{3.70}$$

$$-(S^*S)v + S^*P^*\lambda + \nu S^*v^* = 0 \tag{3.71}$$

Now say

$$R^*u^* = u_1, \quad u_1 \in \mathcal{B}_u^r, \quad S^*v^* = v_1, \quad v_1 \in \mathcal{B}_v^s \tag{3.72}$$

and let by assumption $(R^*R)^{-1}$ and $(S^*S)^{-1}$ exist. Thus we can write the strategies as

$$u^o = -(R^*R)^{-1} R^*L^*\lambda - \mu u_1 \tag{3.73}$$

$$v^o = (S^*S)^{-1} S^*P^*\lambda + \nu v_1 \tag{3.74}$$

and

$$\lambda = Qx \tag{3.75}$$

Thus

$$u^o = -\mu u_1 - (R^*R)^{-1} R^*L^*Qx \tag{3.76}$$

$$v^o = \nu v_1 + (S^*S)^{-1} S^*P^*Qx \tag{3.77}$$

For $\mu = 0$, $\nu = 0$ we must replace (3.76 - 3.77) by the corresponding saturation functions to take care of constraints:

$$u^o = -SAT\ (R^*R)^{-1} R^*L^*Qx \tag{3.78}$$

$$v^o = SAT\ (S^*S)^{-1} S^*P^*Qx \tag{3.79}$$

In the game in section 3.2a we make the identification $x = e$, $x^o = -g$. Then $R^* = R^{-1}$ and $S^* = S^{-1}$ and we have

$$u(t) = -SAT\ R^{-1}(t) \int_t^T B^T(\tau)\ \phi^T(\tau, T)\ Q(t)x(t)dt \tag{3.80}$$

$$v(t) = SAT\ S^{-1}(t) \int_t^T C^T(t)\phi^T(t,T)\ Q(t)\ x(t)dt \tag{3.81}$$

where L^* and P^* are defined as

$$L^*x = \int_t^T B^T(\tau)\ \phi^T(\tau, T)\ x(\tau)d\tau \tag{3.82}$$

$$P^*x = \int_t^T C^T(\tau)\ \phi^T(\tau, T)\ x(\tau)d\tau \tag{3.83}$$

In many nonlinear cases we have to resort to numerical methods in Banach spaces to solve the equations generated by (3.36 - 3.40) to determine the optimal programs. Thus the normal form game solution by this technique is obtained as a pair of optimal programs, while the optimal strategies are required to have a feedback structure.

In the feedback problem or the synthesis problem, the concept of incomplete information to the players must be adequately taken care of. We shall see in the next section how this may be done in the case of partial information. In a purely heuristic way we wish to consider the various problems here. Having determined optimal programs, equivalent strategies based on partial information have to be accounted for in determining feedback strategies. In the process if randomisations or feedback noise enter, the equivalence will no longer be true. If a feedback strategy can be found, it must then satisfy a modification of the corollary to theorem 3.3 which incorporates measures to account for feedback channel noise and randomisations, and here we have to invoke the analog of dual control theory |34| in function spaces.

Supposing we consider the linear game in (3.57 - 3.58) and in addition impose the constraint of partial information to players

$$y = H_1 x \qquad z = H_2 x \tag{3.84}$$

If it is possible to find linear operators Γ_1, Γ_2 such that the players use

$$u = \Gamma_1 y \qquad v = \Gamma_2 z \tag{3.85}$$

we have then a way out. Let

$$u^o = K_1^o x \qquad v^o = K_2^o x \tag{3.86}$$

be the corresponding optimal strategies for the complete information case. We now require the equivalence

$$u^o = K_1^o x = \Gamma_1^o H_1 x \tag{3.87}$$

$$v^o = K_2^o x = \Gamma_2^o H_2 x \tag{3.88}$$

We have now a method of determining Γ_1^o, Γ_2^o,

$$\Gamma_1^o H_1 = K_1^o, \quad \Gamma_2^o H_2 = K_2^o \tag{3.89}$$

Is such an approach possible for other kinds of incomplete information? In the case of sampled and random information, we can define new operators which do perform a similar function as in (3.86).

3.8 GAMES WITH INCOMPLETE INFORMATION: CONTROLLABILITY, OBSERVABILITY AND SUFFICIENT COORDINATES

We consider here only linear differential games, given that the system dynamics are

$$\dot{x} = Ax + Bu + Cv \tag{3.90}$$

where x, u, v are vector functions of time and A, B, C are constant matrices as usual.

(i) The game is of complete information to the players if

$$y = x = z \tag{3.91}$$

(ii) The game is of partial information to the players if

$$y = H_1 x \qquad z = H_2 x \tag{3.92}$$

H_1, H_2 are $m \times n$, $l \times n$ matrices $m \leq n$, $l \leq n$.

(iii) The game is of partitioned information if the observations of players are given as

$$y = \delta_j(x) \qquad z = \delta_k(x) \tag{3.93}$$

where $\delta_j(x)$ and $\delta_k(x)$ are the indicator functions of the sets

$$\left\{ X_j : \quad {}_jX_j = X \,, \quad j = 1,\ldots,\alpha \right\} \tag{3.94}$$

$$\left\{ X_k : \quad {}_kX_k = X \,, \quad k = 1,\ldots,\beta \right\} \tag{3.95}$$

where $\delta_j(x) = 1 \quad x \in X_j \qquad \delta_k(x) = 1 \quad x \in X_k$

$$= 0 \quad x \notin X_j \qquad\qquad = 0 \quad x \notin X_k \tag{3.96}$$

Thus each player knows in which particular set the position lies but is unable to tell its exact magnitude and direction.

(iv) The game is of sampled information if the position is available at discrete instants to the players. As shown in Fig. 3.2 each player has a sampler in his observation equation.

(v) The game is of null information if $y \equiv z \equiv 0$ to the players.

(vi) The game is of random information if

$$\begin{aligned} y &= x + \xi \\ z &= x + \zeta \end{aligned} \tag{3.97}$$

where ξ, ζ are noise processes in measurements.

(vii) The game is of random partial information if

$$\begin{aligned} y &= H_1 x + \xi \\ z &= H_2 x + \zeta \end{aligned} \tag{3.98}$$

We can answer the questions of controllability, observa-bility and sufficient coordinates in cases of complete, partial, sampled and random information. Clearly the problem of random information can also be studied in stochastic games (chapter 5).

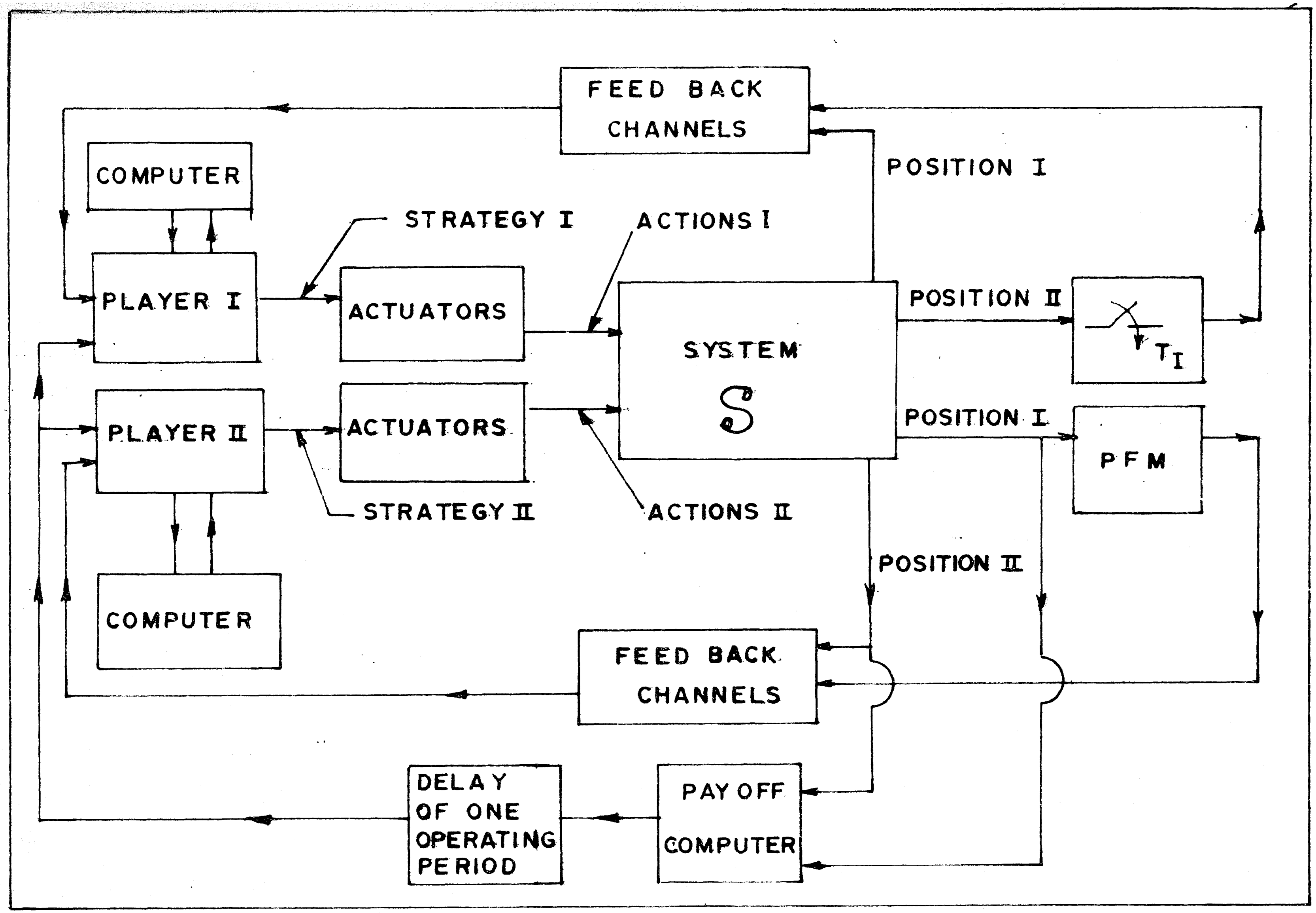


FIGURE 32 A GAME WITH SAMPLED INFORMATION.

We consider here only the complete and partial information case. Let the payoff function be

$$I(u,v) = \langle x(T), Px(T)\rangle + \int_0^T (\langle u(t), Qu(t)\rangle - \langle v(t), Rv(t)\rangle)dt \qquad (3.99)$$

(a) Complete Information: In determining the optimal strategies we need the assurance of a playable pair. A playable pair of strategies in this case turns out as the pair (u,v) such that for every $u\epsilon U$, $v\epsilon V$, the system is controllable. In the game it is binding on both players to start at $x(t_o) = x_o$ specified and terminate the game at $t = T$ on the specified manyfold which in this case is an ϵ-ball, $\|x(T)\| \leq \epsilon$, around the origin $0 < \epsilon < \epsilon_o$, ϵ_o a fixed number. In the limit when $\epsilon_o \to 0$, the controllability condition is precisely that given in [44]. From the second player's viewpoint the system appears as

$$\dot{x} = (A + BK_1(t))x + Cv \qquad (3.100)$$

where he assumes

$$u = K_1(t)x \qquad (3.101)$$

$K_1(t)$ chosen such that $u\epsilon U$. Then for a given K_1 every $v\epsilon V$ must be playable which implies that the pair $(C, (A+BK_1(t)))$ be controllable. In a similar manner consider the game from the viewpoint of player I, to whom the system appears

$$\dot{x} = (A + CK_2(t))x + Bu \qquad (3.102)$$

where he assumes

$$v = K_2(t)x \qquad (3.103)$$

$K_2(t)$ chosen such that $v\epsilon V$. By the same reasoning for a given

K_2 every $u \in U$ must be playable or $(B, (A + CK_2(t)))$ is controllable. In particular, at the optimal point since

$$K_1^*(t) = -Q^{-1} B^T P(t) \tag{3.104}$$

$$K_2^*(t) = R^{-1} C^T P(t) \tag{3.105}$$

and $P(t)$ satisfies the matrix Riccati equation

$$-\dot{P}(t) = P(t)A + A^T P(t) + P(t)\,(CR^{-1}C^T - BQ^{-1}B^T)P(t) \tag{3.106}$$

with

$$P(T) = F \tag{3.107}$$

The matrix $P(t)$ is the same for both players since at the optimal point the adjoint state is unique and $\lambda = P(t)x$. Now let us consider the questions of observability for the game with pe partial information. The above equations can be rewritten as

$$\dot{x} = (A + BK_1^*(t) + CK_2^*(t))x \tag{3.108}$$

$$x(t_0) = x_0 \tag{3.109}$$

$$y = H_1 x \tag{3.110}$$

$$z = H_2 x \tag{3.111}$$

If the system is observable to player I then the matrix

$$\int_{t_0}^{T} \phi^T(T,s)\, H_1^T(s)\, H_1(s)\, \phi(s,t_0)\,ds \tag{3.112}$$

has rank n, where $\phi(t,t_0)$ is the transition matrix given by

$$\frac{d\phi(t,t_0)}{dt} = (A + BK_1^*(t) + CK_2^*(t))\,\phi(t,t_0) \tag{3.113}$$

For player II the observability matrix is replaced by

$$M' = \int_{t_0}^{T} \phi^T(T,s) H_2^T(s) H_2(s) \phi(s,t_0) ds \qquad (3.114)$$

which must have rank n.

The system will be of perfect information if both the matrices have rank n and thus observable to both players. The players can then reconstruct the position vector. The synthesis of the linear feedback laws in terms of $K_1^*(t)$, $K_2^*(t)$ was assumed for the case of partial information. <u>Note that we need to distinguish between partial observability (when the M and M' matrices do not have rank n) and partial information (when not all positions are observed).</u>

(b) <u>Partial Information</u> : If now we require that the optimal strategies be synthesised in terms of the data sets

$$Y_t = \{ y(s) : 0 \leq s \leq t \} \qquad (3.115)$$

$$Z_t = \{ z(s) : 0 \leq s \leq t \} \qquad (3.116)$$

as in the positional game, then let us assume that linear integral operators which map, $\underline{u}: Y_t \to U(t)$, $\underline{v}: Z_t \to V(t)$ exist then we can write

$$u^*(y) = \int_0^t \mathcal{K}_1(t,s) y(s) ds \qquad y(s) \in Y_t \qquad (3.117)$$

$$v^*(z) = \int_0^t \mathcal{K}_1(t,s) z(s) ds \qquad z(s) \in Z_t \qquad (3.118)$$

This requires player I (II) to set up in his computer a storage device of $Y_t (Z_t)$, and an integral operator program to determine $u^*(y)$ ($v^*(z)$). This is rather cumbersome. Player I invokes an implementational supercriterion which says: find a vector $\hat{x}(t)$

such that the following equation holds

$$u^*(y) = J_1(t)\hat{x}(t) = \int_0^t \mathcal{K}_1(t,s)y(s)ds \tag{3.119}$$

where $J_1(t)$ is still to be determined and is such that the original optimal payoff under complete information is unaltered. It is believed that such is possible in games with incomplete information but perfect information since it is possible here to construct an equivalent game with complete information. $\hat{x}(t)$ is then termed a sufficient coordinate vector for the data Y_t. In a similar manner for player II we require

$$v^*(z) = J_2(t)\bar{x}(t) = \int_0^t \mathcal{K}_2(t,s)z(s)ds \tag{3.120}$$

$\bar{x}(t)$ is the sufficient coordinate vector for Z_t. It turns out that the coordinates $\hat{x}(t)$, $\bar{x}(t)$ are nothing but the respective outputs of reconstructing observer systems, x_0 being known to both players. We shall demonstrate the reconstructing observer for player I. As player I views the game he observes the system

$$\dot{x}(t) = (A + CK_2^*(t))x(t) + Bu(t) \tag{3.121}$$

$$y(t) = H_1x(t) \tag{3.122}$$

The assumption is valid since player I knows under perfect information that player II has to employ $v = K_2^*(t)x$. [In the random information case it is better to write $v = K_2^*(t)\bar{x}$, where $\bar{x}$ is the estimate of x for player II. Since the above game has no random information $x = \bar{x} = \hat{x}$.] Let $S(t)$ be a nonsingular matrix for all t such that

$$\begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix} = w(t) = S(t)\, x(t) \tag{3.123}$$

Then

$$\dot{w}(t) = S^{-1}(t)\,(A + GK_2^*(t))S(t)x(t) + S^{-1}(t)Bu(t) \qquad (3.124)$$

and choose $S(t)$ such that

$$y(t) = H_1 S^{-1}(t)w(t) = [\,I \quad 0\,] \begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix} \qquad (3.125)$$

Now consider the system

$$\begin{bmatrix} \dot{w}_1(t) \\ \dot{w}_2(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} u(t) \qquad (3.126)$$

$$y(t) = w_1(t) \qquad (3.127)$$

Then w_2 can be simply reconstructed by the redundant equation

$$\dot{w}_2 = a_{21}w_1 + a_{22}w_2 + b_2 u \qquad (3.128)$$

$$= a_{21}y + a_{22}w_2 + b_2 u \qquad (3.129)$$

Then the reconstructed state is

$$x = S^{-1}(t) \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = S^{-1}(t) \begin{bmatrix} y \\ w_2 \end{bmatrix} = x \qquad (3.130)$$

and hence the optimal strategy for player I is

$$u^* = K_1^*(t)x \qquad (3.131)$$

and the above is possible since $x(t_o)$ is known and hence $w(t_o)$ is known. In a similar manner player II can have a reconstructing observer for the system

$$\dot{x} = (\,A + BK_1^*(t)x\,)x + Gv$$

$$z = H_2 x. \qquad (3.132)$$

The reconstructing observer is shown in Fig. 3.3 for player I. We shall consider linear games with random information in chapter 5.

3.6 CONCLUSIONS

Function space methods are useful for considering many theoretical aspects for games in normal form. The search for operators that map infinite dimensional spaces into finite dimensional spaces are easy to visualise conceptually. With these, a reduction in computational effort is possible. For problems involving memory and information constraints, however, this method is not suitable.

For linear games we have shown the importance of the concepts of controllability, observability and sufficient coordinates. These concepts can be extended to the N-person game considered in the next chapter.

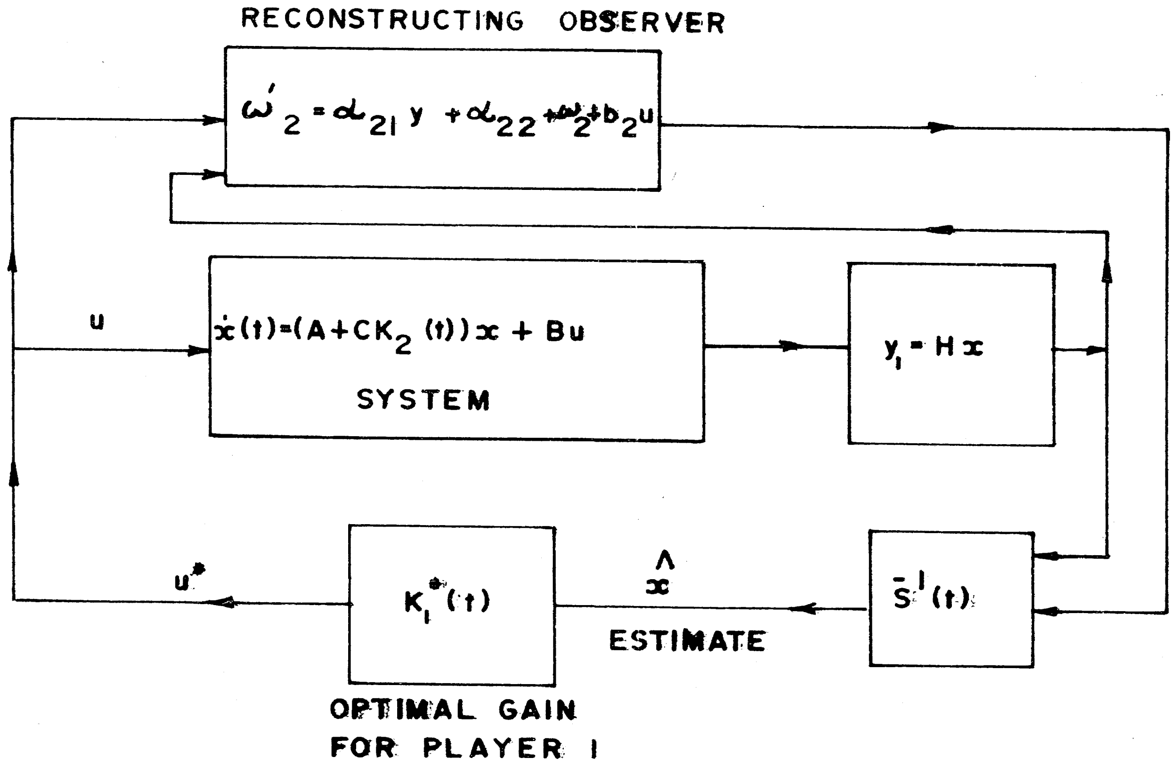


FIGURE 33 RECONSTRUCTING OBSERVER FOR PLAYER I

IV N-PERSON DIFFERENTIAL GAMES

N-person differential games were first constructed by Petrosyan |17| and a maximum principle thereof for workable purposes was given by Kuznetsov and Karvovskiy |18|. In |45| a rigorous derivation of the theory has been given using a dynamic programming approach for the non-cooperative game. References |46-47| also contain more about N-person differential games and the theory for the cooperative versions will be reported in a forthcoming thesis due to Prasad |40|.

When N-decision makers or players have an influence on the outcome of the system, they are playing a game, each perhaps with a different objective. The problems that arise due to the existence of the conflict of goals of N-rational players is certainly a question which cannot in the strict sense be answered either within the confines of engineering or mathematics. Many concepts from the social sciences are required and shall be used |40|. In control systems engineering, an N-person game appears in different guises. We shall see a few examples at the end of this chapter.

In section 2 we give a heuristic derivation of the Minimum principle for N-players. In section 3 we solve a few N-person games and formulate others.

4.2 THE NON-COOPERATIVE THEORY OF N-PERSON DIFFERENTIAL GAMES

We consider here the game $(X,\underline{U},\underline{Y},\delta,\underline{H},e,\underline{I})$ with $\underline{Y} \triangleq (Y^1,Y^2,\ldots,Y^N)$, $Y^i = X$ for all $i = 1,\ldots,N$, $\underline{U} \triangleq (U^1,U^2,\ldots,U^N)$, $\underline{H} \triangleq (H_1,H_2,\ldots,H_N)$, $\underline{I} \triangleq (I^1,I^2,\ldots,I^N)$. Let $x \in X$ be an

n-dimensional vector $u^i \in U^i$ be an r^i-dimensional vector, $i = 1,\dots,N$, $\Theta = [0,T]$. Let $\mathcal{S} = X \times \underline{U} \times \Theta$, $\mathcal{G} = X \times \Theta$. The mapping S induces the function G on $\mathcal{G}$, $\dot{G} \subseteq R^n$,

$$\dot{x} = G(x,\underline{u},t) \tag{4.1}$$

The payoff function to each player is determined by

$$I^i(\underline{u}) = \int_0^T f^i(x,\underline{u},t)dt + g^i(\, x(T),T \,) \tag{4.2}$$

The functions f^i, g^i, G are assumed to be differentiable at least once. Further the rules of the game prescribe certain input constraints on the N-players of the form

$$K^i(x,u^i,t) \geq 0 \tag{4.3}$$

for $i = 1,\dots,N$, and that the game shall start at a prespecified initial position $x_0 \in X$ and shall end on a terminal surface given parametrically as

$$x = x(\sigma) \qquad T = T(\sigma) \tag{4.4}$$

where $\sigma = (\sigma_1,\dots,\sigma_n)$ is a vector of n free parameters, f^i is termed the incremental payoff of player i and g_i the terminal payoff.

We require no randomizations in the policies, hence no set W_q need be included above. We also consider the system with no distributed or stochastic parameters. Hence, out of all possible control policies we should pick out those that assure termination of the game and remain pure. We describe this procedure. Since it is possible that a player may choose to switch his strategy or be faced with a dilemma to choose, at some point $(x,t) \in \mathcal{G}$, we set out to demarcate all possible such points in the $\mathcal{G}$

space that have a special behaviour about them. We expect, however, that these points do not fill the entire Θ space, and lie only on some surfaces or manifolds. Thus the entire Θ space can be divided into different regions. In the interior of each such region only a unique N-tuple of strategy is optimal. Such a decomposition has been rigorously considered by Berkovitz |50|, and we follow the same decomposition.

Let us consider an optimal trajectory. It has to cross many such manifolds before it terminates on Σ, the terminal surface. Let (ξ, τ) be in the interior of one such region lying on the optimal trajectory as shown in Fig. 4.1. Consider now the set of all control actions of all players $\underline{U}$. For $\underline{u} \in \underline{U}$ it is possible to find some x_o from which the system can be started and the game terminated at Σ. If we change one u^i, then this may no longer be true. Consider now a subset $\underline{U}_o \subseteq \underline{U}$ such that any u^i and u^j in $\underline{U}_o$ are admissible in the sense that they transfer of the system from x_o to Σ is assured. $\underline{u} \in \underline{U}_o$ is then said to be a playable N-tuple. Associated with every playable N-tuple, the payoff is single valued to a player. We now define

<u>Definition 4.1</u> : A Nash equilibrium point $\underline{u}^*$ for the N-person game relative to strategies $\underline{U}_o$ is said to exist if

$$I^i(x_o, (\underline{u}^*; u^i), t_o) \geq I^i(x_o, \underline{u}^*, t_o) \tag{4.8}$$

for every $i = 1, \ldots, N$ where the notation $(\underline{u}^*; u^i)$ stands for u^i not being optimal when the rest are. (Refer |49|).

We shall now proceed to determine the conditions under which the game can have a Nash equilibrium point. In doing so

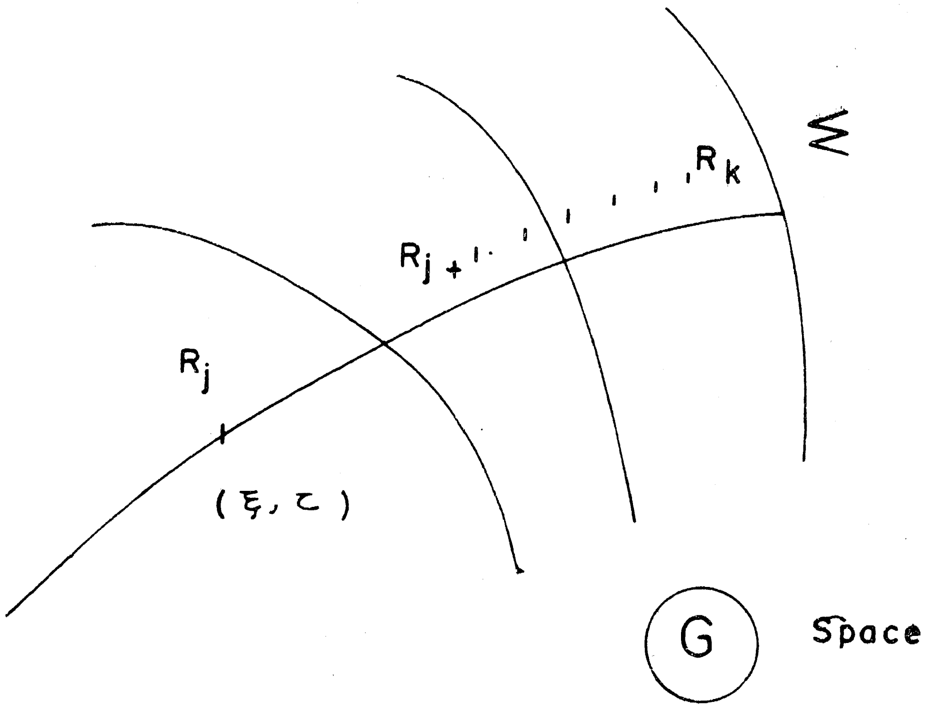


FIGURE 4·1 THE DECOMPOSITION OF THE (G) SPACE ALONG THE TRAJECTORY.

we have the following conditions to bear in mind.

(i) The decomposition associated with $\underline{\mu}^*$ is regular.

(ii) If the initial point is interior to one of the regions then the optimal trajectory is unique till it hits a manifold of discontinuity. If it is a switching surface, it will continue into the next region. If it is any one of the surfaces discussed by Isaacs, |1|, the corner conditions have to be appropriately determined.

(iii) The optimal path is never tangent to an interior manifold or a terminal manifold.

(iv) We consider the existence of a unique equilibrium point or perfect knowledge among players as to which equilibrium point strategies they would play.

Only necessary conditions for an N-tuple of strategies to be an equilibrium point N-tuple of strategies are determined.

(a) <u>The Hamilton-Jacobi-Bellman equations</u> : We now determine the Hamilton-Jacobi-Bellman partial differential equations and the minimax principle as satisfied by each player. From the point of view of each player, the determination of the optimal payoff under the assumption of optimality of the strategies of the other players leads to an associated optimal control problem with the differential constraint in (4.1).

Let $V^i(\xi,\tau)$ denote the optimal payoff to player i, starting the game from (ξ,τ) and using the optimal strategy u^{i*}.

$$V^i(\xi,\tau) = g^{*i}(x(T),T) + \int_\tau^{t_j} f^{i*}dt + \Sigma_{l=j+1}^{k} \int_{t_{l-1}}^{t_l} f^{i*}dt \quad (4.6)$$

where $$f^{*i} = f^i(x^*, \underline{u}^*, t) \tag{4.7}$$

and k depends on the decomposition. We assume that the derivative W^i_ξ, W^i_τ exist for all i, ξ, τ, and these are continuous through G. To determine the Nash equilibrium point we assume a particular nonoptimal strategy for the ith player.

$$\bar{u}^i(x,t) = \begin{cases} u^{*i}(x,t) & (x,t) \notin N(\xi,\tau) \\ u(x,t) & (x,t) \in N(\xi,\tau) \end{cases} \tag{4.8}$$

where $N(\xi,\tau)$ is a δ-neighbourhood of (ξ,τ) so chosen that it is wholly within R_j, and this strategy is playable against all others in $\underline{U}_0$. δ stands for the last time the trajectory leaves the neighbourhood $N(\xi,\tau)$. Now

$$W^i(\xi,\tau) = I^i(\xi, \underline{u}^*, \tau) \leq I^i(\xi, (\underline{u}^*; u^i), \tau) \tag{4.9}$$

The right hand side of the inequality can be expanded as

$$g^i(x(T),T) + \left(\int_\tau^{\tau+\delta} + \int_{\tau+\delta}^T\right) f^i(x(t), (\underline{u}^*(\xi,\tau); \bar{u}^i(\xi,\tau)), t)\,dt$$

$$= \int_\tau^{\tau+\delta} f^i(x(t), (\underline{u}^*(x(t),t); \bar{u}^i(x(t),t)), t)\,dt + W^i(x(\ +\delta),\ +\delta) \tag{4.10}$$

Hence

$$-W^i(x(\tau+\delta), \tau+\delta) + W^i(\xi,\tau)$$
$$\leq \int_\tau^{\tau+\delta} f^i(x(t), (\underline{u}^*(x(t),t); \bar{u}^i(x(t),t)), t)\,dt \tag{4.11}$$

Now if we let $\delta \to 0$, $x(\tau+\delta) \to \xi$, then we can write the left hand side of the above as

$$-W^i_\tau(\xi,\tau)\delta - W^i_\xi(\xi,\tau)\,[\,x(\tau+\delta) - \xi\,] + o(\delta)$$

$$= -W^i_\tau(\xi,\tau)\delta - W^i_\xi(\xi,\tau)G(\xi,\ (\underline{u}^*(\xi,\tau);\bar{u}^i(\xi,\tau)\ ),\tau)\delta + o(\delta) \tag{4.12}$$

The right hand side of (4.11) can be written by the mean value theorem as $f^i(\xi,\ (\underline{u}^*(\xi,\tau);\ \bar{u}^i(\xi,\tau)\ ),\tau)\delta$ which finally leads us to write

$$-W^i_\tau(\xi,\tau) \leq f^i(\xi,(\underline{u}^*(\xi,\tau);\ \bar{u}^i(\xi,\tau)\ ),\tau\ )$$

$$+\ W^i_\xi(\xi,\tau)\ G(\xi,\ (\underline{u}^*(\xi,\tau);\ \bar{u}^i(\xi,\tau)\ ),\tau) \tag{4.13}$$

Since $u^i \in U^i_o$ is arbitrary, (4.13) holds for all u^i in U^i_o. The equality holds for $u^i = u^{i*}$. The inequality holds for each $i=1,\dots,N$. The above constitutes necessary condition of a Nash equilibrium point tuple of strategies. Alternatively consider the N-person game with payoff to each player

$$f^i(\ x,\underline{u}(x,t),t\ ) + W^i_x(x,t)\ G(\ x,\underline{u}(x,t),t\ ) \tag{4.14}$$

For the condition where (ξ,τ) lies on any M_j the continuity of W_ξ, W_τ can be established if M_j is a manifold of discontinuity of all but one player. When it is a manifold of discontinuity of all players a corner condition holds as in the 2-person game theory. Let us now introduce the functions

$$H^i(x,\underline{u},\lambda^i,t) = f^i(x,\underline{u},t) + \lambda^i G(x,\underline{u},t) \tag{4.15}$$

which we term as the Hamiltonian for player i. Then for each Hamiltonian we require

$$H^i(x,\underline{u}^*(x,t),\ W^i_x(x,t),t) + W^i_t(x,t) = 0 \tag{4.16}$$

We can give an alternative form for the necessary conditions in terms of the Hamiltonians for the players.

(b) <u>The Hamiltonian Form or the Minimum Principle for the Players</u> :

We now determine the Hamiltonian form of |18| from (4.15). This is obtained by relating W^i_x and W^i_t to the adjoint variables. For simplicity consider only the region R_j immediately preceding the terminal surface Σ_j. Let $(\xi,\tau)\in R_j$. Then

$$W^i(\xi,\tau) = \frac{\partial K^i(\overset{*}{x}(T),T)}{\partial x(T)}\frac{\partial \overset{*}{x}(T)}{\partial\sigma}\frac{\partial\sigma}{\partial\xi} + f^i(\overset{*}{x}(T),T)\frac{\partial T}{\partial\sigma}$$

$$+ \frac{\partial K^i(\overset{*}{x}(T),T)}{\partial T}\frac{\partial T}{\partial\sigma}\frac{\partial\sigma}{\partial\xi}$$

$$+ \int_\tau^T \left(\frac{\partial f^{i*}}{\partial x}(x^*(t),t) + \Sigma_{l=1}^{N}\frac{\partial f^{i*}}{\partial u^{l*}}\frac{\partial u^{l*}(x^*(t),t)}{\partial x}\frac{\partial x^*}{\partial\sigma}(t)\right)$$

(4.17)

We shall simplify this expression through equation (4.15 - 4.16). Consider

$$\frac{\partial K^i(x^*(T),T)}{\partial x(T)}\frac{\partial \overset{*}{x}(T)}{\partial\sigma} + \frac{\partial K^i(x^*(T),T)}{\partial T}\frac{\partial T}{\partial\sigma} + f^i(x^*(T),T)\frac{\partial T}{\partial\sigma}$$

$$+ \lambda^i(T)\left(g\left(x^*(T), \underline{u}^*(x^*(T),T)\frac{\partial T}{\partial\sigma} - \frac{\partial x^*(T)}{\partial\sigma}\right)\right) = 0 \quad (4.18)$$

As assumed $\sigma\in\mathcal{R}^n$, hence if the $n \times n$ coefficient matrix of $\lambda^i(T)$ is nonsingular, $\lambda^i(T)$ is uniquely defined and (4.18) can be simplified to read

$$\frac{\partial K^i(x^*(T),T)}{\partial x^*(T)}\frac{\partial x^*(T)}{\partial\sigma}\frac{\partial\sigma}{\partial\xi} + \frac{\partial K^i(x^*(T),T)}{\partial T}\frac{\partial T}{\partial\sigma}\frac{\partial\sigma}{\partial\xi}$$

$$+ \left(H^i\frac{\partial T}{\partial\sigma}\frac{\partial\sigma}{\partial\xi} - \lambda^i(T)\frac{\partial x^*(T)}{\partial\sigma}\frac{\partial\sigma}{\partial\xi}\right) = 0 \quad (4.19)$$

or if $\underline{M}$ stands for the coefficient matrix we have

$$\lambda^i(T) = \frac{\partial G^i(x^*(T),T)}{\partial x^*(T)} \frac{\partial x^*(T)}{\partial \sigma} + \frac{\partial G^i(x^*(T),T)}{\partial T} \frac{\partial T}{\partial \sigma}$$

$$+ f^i(x^*(T),T) \frac{\partial T}{\partial \sigma} \Big\} \quad M^{-1} \tag{4.20}$$

Now consider the system of linear differential equations

$$\frac{d\lambda^i}{dt} = -\frac{\partial H^{*i}}{\partial x} - \Sigma_{i=1}^{N} \frac{\partial H^{*i}}{\partial u^i} \frac{\partial u^{i*}}{\partial x^*} \tag{4.21}$$

with $\lambda^i(T)$ specified in (4.20). A unique solution to this exists and hence as long as the point (ξ, τ) is interior to any region, λ^i is continuous. It can be shown to remain continuous across all manifolds of discontinuity where all but one player switches his strategy. We shall see now how the constraint condition in (4.3) can be incorporated.

Let $K^i(x,u^i,t)$ be a piecewise continuous function and let $K^i \in E^{p^i}$. Then if $p^i > r^i$ at each point of (x,u^i,t) at most r^i of the K^i can vanish. The matrix $\left\{\frac{\partial K_k^i}{\partial u_l^i}\right\}$, $k = 1,\ldots,p^i$, $l = 1,\ldots,r^i$, has maximum rank. Let $\mu^i(x,t)$ be multiplier functions such that

$$\frac{\partial H^i}{\partial u^{i*}} + \mu^i \frac{\partial K^i}{\partial u^i} = 0 \tag{4.22}$$

$$\mu_j^i K_j^i = 0, \quad j = 1,\ldots,r^i, \ \mu^i \leq 0 \tag{4.23}$$

for every $i = 1,\ldots,N$. Then equation (4.21) can be written as

$$\frac{d\lambda^i}{dt} = -\frac{\partial H^{*i}}{\partial x} - \Sigma_{i=1}^{N} \mu^i \frac{\partial K^i}{\partial x} \tag{4.24}$$

This easily follows since from (4.22) we have

$$-\mu^i \frac{\partial K^i}{\partial u^i} = \frac{\partial H^i}{\partial u^{i*}} \tag{4.25}$$

At the optimal point each component of $K^i(x,u^i,t)$ is either zero or a relative minimum. Hence we must have on substituting $u^i = u^{i*}(x^*(t),t)$, and

$$\frac{\partial K^i}{\partial x^*} + \frac{\partial K^i}{\partial u^i} \frac{\partial u^{i*}}{\partial x^*} = 0 \tag{4.26}$$

Hence

$$\mu^i \frac{\partial K^i}{\partial x^*} + \mu^i \frac{\partial K^i}{\partial u^i} \frac{\partial u^{i*}}{\partial x^*} = 0 \tag{4.27}$$

From (4.25) and (4.27) it follows that

$$\frac{\partial H^{i*}}{\partial u^i} \quad \frac{\partial u^{i*}}{\partial x^*} = \mu^i \frac{\partial K^i}{\partial x^*} \tag{4.28}$$

In terms of the N-Hamiltonians we can rewrite a stronger condition from (4.15) and (4.16)

$$H^i(x^*, (\underline{u}^*,u^i),t) \geq H^i(x^*,\underline{u}^*,t) \tag{4.29}$$

for each $i=1,\ldots,N$.

We have thus generalized the result given in [18]. As a corollary, the results for two-person zero-sum games can be obtained by setting $N=2$, $I^1(u^1,u^2) = -I^2(u^1,u^2)$. Perhaps, we could also consider the relaxation of the requirement of n-dimensionality of the terminal surface.

4.3 EXAMPLES

In this section we mainly study some pedagogical N-person games in systems engineering as an application of the theory in section 4.2. Our first example is from an economic context.

<u>Example 1</u> : <u>The Non-cooperative Game of the Economy of Two Nations</u> : There are two nations, 1 and 2, whose state of economy is given by the cumulative growth of its vital resources vector and its total production vector. The totality of these is depicted by the vectors X^1 and X^2 for 1 and 2 respectively. Each nation has to make a minimal contribution to each other's economy.

$B_1^1X^1$ to the economy of 1 of X^1

$B_1^2X^2$ to the economy of 1 of X^2

$B_2^1X^1$ to the economy of 2 of X^1

$B_2^2X^2$ to the economy of 2 of X^2

where B_i^j is the matrix of minimal per unit allocation of resources j to economy i. The societies of each nation have a requirement of minimal growth rate given by D^1 and D^2. In addition nation i can make controlled contributions to the economy of j, this is given by the matrix $(A_j^iU_j^i)$ which consists of controlled per unit additional allocations of resources by nation i to nation j. The fractions U_j^i have constraints

$$\Sigma_{k=1}^{2} \Sigma_{l=1}^{n_k} \Sigma_{q=1}^{n_i} U_{kl}^{iq} \leq 1 \qquad i,k = 1,2 \tag{4.30}$$

and the goals of the economies are under a finite horizon plan at the end of which each nation j wants

$$g^j = C^jX^j \tag{4.31}$$

to be maximum subject to

$$X^j(t) \geq 0 \qquad \text{for all} \quad t \in [0,T] \tag{4.32}$$

where $[0,T]$ is the duration of the plan. The dynamical growth is given by

$$\frac{dX^1}{dt} = (A_1^1U_1^1)\,X^1 + (A_1^2U_1^2)X^2 + B_1^1X^1 + B_1^2X^2 + D^1 \tag{4.33}$$

$$\frac{dX^2}{dt} = (A_2^1U_2^1)\,X^1 + (A_2^2U_2^2)X^2 + B_2^1X^1 + B_2^2X^2 + D^2 \tag{4.34}$$

A block diagram model of the economy is shown in Fig. 4.2. Let the vectors λ_1^1 and λ_1^2 be the shadow prices (incremental) as seen by nation 1 of the economies 1 and 2 respectively. Let λ_1^2 and λ_2^2 be the incremental shadow price vectors as seen by nation 2 of the economies of 1 and 2 respectively. To obtain the optimum policies under non-cooperation, consider the Hamiltonians

$$H^1 = \lambda_1^1\,[\,(A_1^1U_1^1)\,X^1 + (A_1^2U_1^2)X^2 + B_1^1X^1 + B_1^2X^2 + D^1\,]$$

$$+ \lambda_2^1\,[\,(A_2^1U_2^1)\,X^1 + (A_2^2U_2^2)X^2 + B_2^1X^1 + B_2^2X^2 + D^2\,] \tag{4.35}$$

$$H^2 = \lambda_1^2\,[\,(A_1^1U_1^1)X^1 + (A_1^2U_1^2)X^2 + B_1^1X_1^1 + B_1^2X^2 + D^1]$$

$$+ \lambda_2^2\,[\,(A_2^1U_2^1)X^1 + (A_2^2U_2^2)X^2 + B_2^1X^1 + B_2^2X^2 + D^2\,] \tag{4.36}$$

We have to maximise H^1 for choices of U^1 and H^2 for choices of U^2. This is obtained by considering only

$$\max_{U^1 \in \underline{R}^1} H_1^1 = \lambda_1^1[\,(A_1^1U_1^1)X^1\,] + \lambda_2^1[(A_2^1U_2^1)X^1] \tag{4.37}$$

where $\underline{R}^1$ is the projection of the constraints on the U^1 space. Similarly

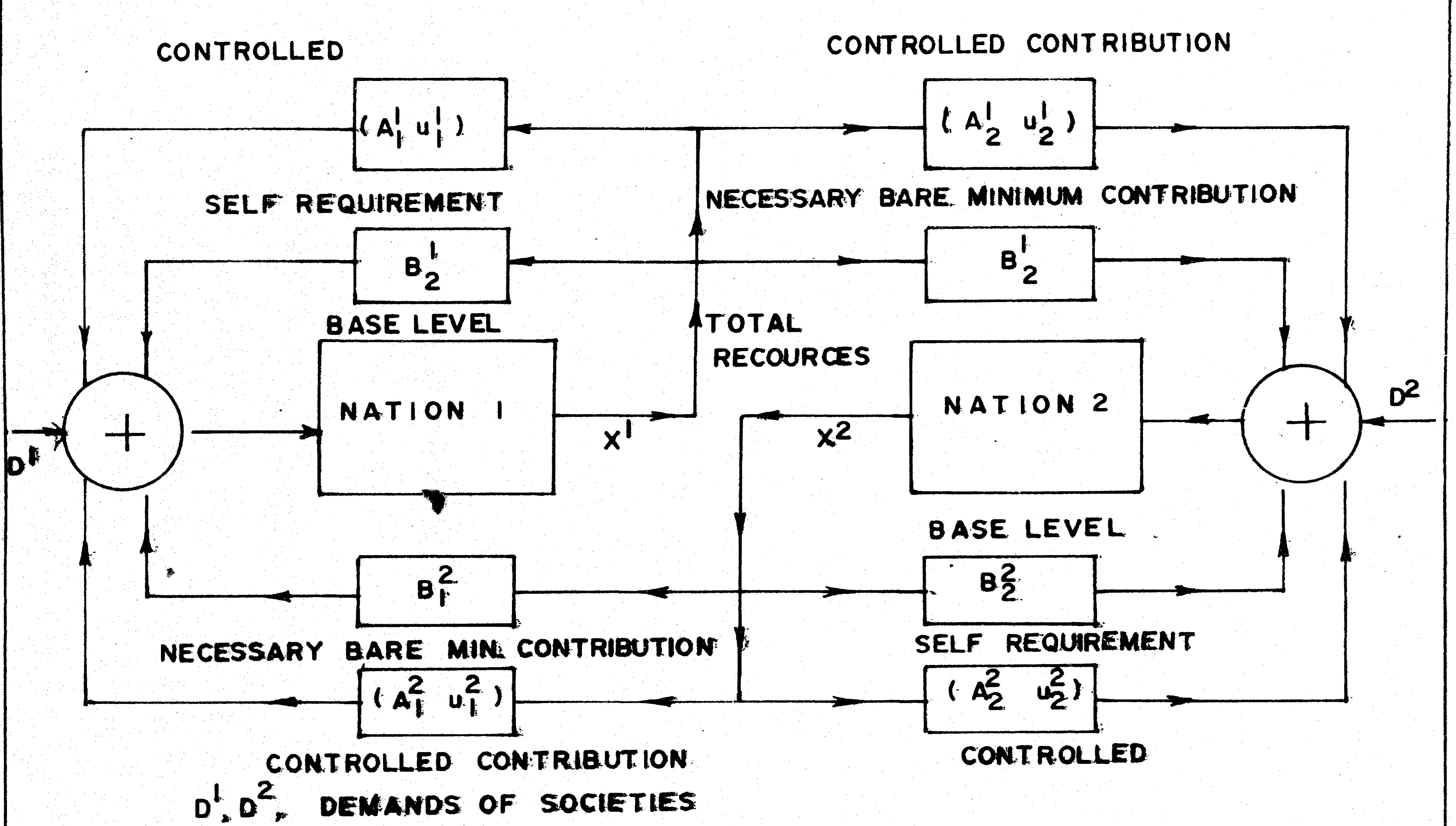


FIGURE 4.2 THE NATIONAL ECONOMY MODEL OF TWO NATIONS.

$$\max_{U^2 \in R^2} H_2^2 = \lambda_1^2[(A_1^2 U_1^2) x^2] + \lambda_2^2[(A_2^2 U_2^2) x^2] \tag{4.38}$$

The adjoint equations are

$$\begin{aligned}
\dot\lambda_1^1 &= -\lambda_1^1[(A_1^1 U_1^1) + B_1^1)] - \lambda_2^1[(A_2^1 U_2^1) + B_2^1] \\
\dot\lambda_2^1 &= -\lambda_1^1[(A_1^2 U_1^2) + B_1^2)] - \lambda_2^1[(A_2^2 U_2^2) + B_2^2] \\
\dot\lambda_1^2 &= -\lambda_1^2[(A_1^1 U_1^1) + B_1^1] - \lambda_2^2[(A_2^1 U_2^1) + B_2^1] \\
\dot\lambda_2^2 &= -\lambda_1^2[(A_1^2 U_1^2) + B_1^2] - \lambda_2^2[(A_2^2 U_2^2) + B_2^2]
\end{aligned} \tag{4.39}$$

with terminal conditions

$$\begin{aligned}
\lambda_1^1 &= C^1 & \lambda_2^1 &= 0 \\
\lambda_1^2 &= 0 & \lambda_2^2 &= C^2
\end{aligned} \tag{4.40}$$

If the nations cooperate, their economies are bound to prosper. Such is considered by Freund |48|.

<u>Example 2</u> : <u>A Bicriterion Control Problem</u> : One potential application of N-person game theory is in problems of system optimisation with multivalued criteria. The example chosen here represents the situation where there exists a dilemma whether to design a system time-optimally or fuel-optimally. We settle for a compromise optimal design for both criteria in the sense defined as follows. This is done by allocating the control input between two constituent 'inputs'. One 'input' is now chosen according to time-optimality, the other is chosen for fuel-optimality. [In

economic terms, would we call them minimum time (Harrison) plans and minimum aid plans?] Perhaps, this is also representative of the hypothetical situation where with limited resources the guidance of a vehicle under the command of several 'pilots' leads to different criteria being employed by each one of them for the same goal.

We study the following single axis satellite attitude control problem by reformulating it as a differential game. We are interested in finding the behaviour of the satellite under the influence of its gas jet controllers with limited thrust. The terminal target is to reach the origin of the error and error-rate plane. Let

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= c_1u_1 + c_2u_2 \end{aligned} \tag{4.41}$$

be the dynamical description of the vehicle under consideration, where x_1 is the attitude error-angle, x_2 is the attitude error-rate angle, c_1 and c_2 are the coefficients influenceable by the designer, while u_1, u_2 are the control strategies to be chosen for different criteria by treating them as the respective control strategies of two different players. Their objective functions are

$$I^1(u_1) = \int_0^t dt \tag{4.42}$$

$$I^2(u_2) = \int_0^t |u_2|\,dt \tag{4.43}$$

The corresponding Hamiltonians are

$$H^1 = 1 + \lambda_1^1 x + \lambda_2^1(c_1u_1 + c_2u_2) \tag{4.44}$$

$$H^2 = |u_2| + \lambda_1^2 x_2 + \lambda_2^2(c_1u_1 + c_2u_2) \tag{4.45}$$

The optimality conditions state that H^1 should be minimised for u_1 and H^2 for u_2. Thus

$$u_1^* = -\text{sign}\,\lambda_2^1 c_1 \tag{4.46}$$

$$\begin{aligned} u_2^* &= -\text{sign}\,\lambda_2^2 c_2 && \text{if } |\lambda_2^2 c_2| > 1 \\ &= -v(t)\,\text{sign}\,\lambda_2^2 c_2 && \text{if } |\lambda_2^2 c_2| = 1, 0 \leq |v(t)| \leq 1 \\ &= 0 && \text{if } |\lambda_2^2 c_2| \leq 1 \end{aligned} \tag{4.47}$$

Let us choose the surface $x_1 = r\cos\theta$, $x_2 = r\sin\theta$, with r, θ as free variable. Then

$$[\lambda_1^1(T) \quad \lambda_2^1(T)]\begin{bmatrix} r\sin\theta & -r\sin\theta \\ c_1u_1 + c_2u_2 & -r\cos\theta \end{bmatrix} = -[1 \quad 0] \tag{4.48}$$

$$\lambda_1^1(T) = \lambda_2^1(T)\cot\theta \qquad \lambda_2^1(T) = \frac{-1}{r\cos\theta + c_1u_1 + c_2u_2} \tag{4.49}$$

$$[\lambda_1^2(T) \quad \lambda_2^2(T)]\begin{bmatrix} r\sin\theta & -r\sin\theta \\ c_1u_1 + c_2u_2 & r\cos\theta \end{bmatrix} = -[1 \quad 0]\,|u_2| \tag{4.50}$$

$$\lambda_1^2(T) = \lambda_2^2(T)\cot\theta \qquad \lambda_2^2(T) = \frac{-|u_2|}{r\cos\theta + c_1u_1 + c_2u_2} \tag{4.51}$$

$$\frac{\lambda_2^2(T)}{\lambda_2^1(T)} = |u_2| \tag{4.52}$$

We consider the limiting situation as $r \to 0$. If $|u_2| = 1$, then

$$\lambda_2^2(T) = \lambda_2^1(T) = \frac{\pm 1}{c_1 + c_2}, \quad \lambda_1^1(T) = \lambda_1^2(T) = a\,(\text{const.say}) \tag{4.53}$$

Then $\lambda_2^2(t) = \lambda_2^1(t) = \dfrac{-\text{sign } u_1}{c_1 + c_2} + \alpha(T-t), \; t < T$ (4.54)

If $|u_2| = \epsilon < 1$ then $\lambda_2^1(T) = \dfrac{-\text{sign } u_1}{c_1 + \epsilon c_2}$

(4.55)

$$\lambda_2^2(T) = \frac{-\epsilon \text{ sign } u_1}{c_1 + \epsilon c_2}$$

$$\lambda_1^1(T) = \alpha \text{ (const. say)}, \quad \lambda_1^2(T) = \epsilon\alpha \qquad (4.56)$$

Hence $\lambda_2^2(t) = \dfrac{-\text{sign } u_1}{c_1 + \epsilon c_2} + \alpha\,(T-t) \quad t < T$ (4.57)

$$\lambda_2^1(t) = \frac{-\epsilon \text{ sign } u_1}{c_1 + \epsilon c_2} + \epsilon\alpha(T-t) \quad t < T \qquad (4.58)$$

If $|u_2| = 0$, then $\lambda_2^1(T) = \dfrac{-\text{sign } u_1}{c_1}$

(4.59)

$$\lambda_1^2(T) = 0, \lambda_2^2(T) = 0, \lambda_1^1(T) = \alpha \text{(const. say)}$$

$$\lambda_1^1(t) = \frac{-\text{sign } u_1}{c_1} + \alpha(T-t) \quad t < T \qquad (4.60)$$

$$\lambda_1^2(t) = \lambda_1^2(T) = 0 \qquad (4.61)$$

The various possibilities for $\lambda_1^1(t), \lambda_2^1(t), \lambda_1^2(t), \lambda_2^2(t)$ are shown in Fig. 4.3. Hence the control strategy sequences to reach the origin can be obtained from equations (4.53 - 4.61). These and other results are given in the form of lemmas.

<u>Lemma 4.1</u> : The feasible control strategy sequences to reach the origin without a switching are

$$\begin{vmatrix} u_1^0 \\ u_2^0 \end{vmatrix} = \begin{vmatrix} 1 \\ 1 \end{vmatrix}, \begin{vmatrix} -1 \\ 0 \end{vmatrix}, \begin{vmatrix} -1 \\ -1 \end{vmatrix}, \begin{vmatrix} 1 \\ 0 \end{vmatrix} \qquad (4.62)$$

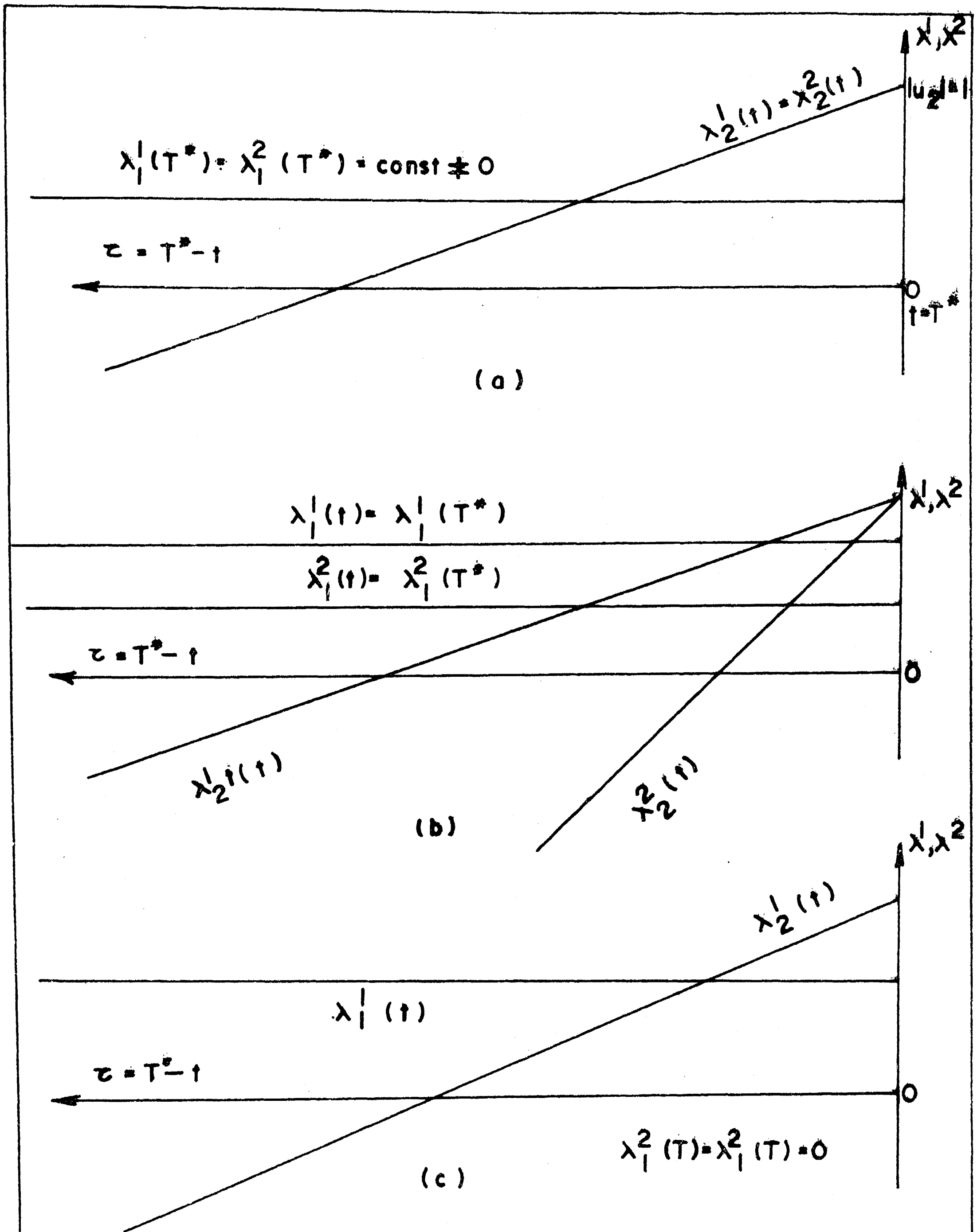


FIGURE 4.3 ADJOINT STATE-TRAJECTORIES FOR THE CASES (a) $|u_2|=1$, (b) $|u_2|<1$ (c) $(u_2)=0$ AT TIME T^*.

Proof : We prove this by substituting (4.62) into (4.41) and then integrating,

$$\dot{x}_1 = x_2 \qquad \dot{x}_2 = \pm (c_1 + c_2) \text{ or } \pm\, c_1 . \qquad \text{Let}$$

$x_1(0) = \beta_1 \quad x_2(0) = \beta_2$. Then

$$x_2 = \pm (c_1 + c_2)t + \beta_2 \quad \text{or} \quad \pm\, c_1 t + \beta_2 . \qquad \text{For}$$

$x_2(T) = 0, \; x_1(T) = 0$ we require

$$x_2 = -(\text{sign } \beta_2)\,(c_1 + c_2)t + \beta_2 \quad \text{or} \quad -(\text{sign } \beta_2)c_1 t + \beta_2$$

$$x_1 = -(\text{sign } \beta_2)\,(c_1 + c_2)\frac{t^2}{2} + \beta_2 t + \beta_1 \quad \text{or} \quad -(\text{sign } \beta_2)c_1\frac{t^2}{2} + \beta_2 t + \beta_1$$

In either case there exists a $T > 0$ such that at $t = T$, $x_1(T) = 0, \; x_2(T) = 0$ where

$$T = \frac{\beta_2}{(\text{sign } \beta_2)(c_1+c_2)} \quad \text{or} \quad \frac{\beta_2}{(\text{sign } \beta_2)\, c_1} \tag{4.63}$$

Q.E.D.

Let

$$\begin{aligned}
\gamma_1^+ &= \left\{ (x_1, x_2) : x_1 = \frac{1}{2c_1} x_2^2 \right\} \\
\gamma_2^+ &= \left\{ (x_1, x_2) : x_1 = \frac{1}{2(c_1+c_2)} x_2^2 \right\} \\
\gamma_1^- &= \left\{ (x_1, x_2) : x_1 = \frac{-1}{2c_1} x_2^2 \right\} \\
\gamma_2^- &= \left\{ (x_1, x_2) : x_1 = \frac{-1}{2(c_1+c_2)} x_2^2 \right\}
\end{aligned} \tag{4.64}$$

Then the truth of the following lemma is easily verified.

<u>Lemma 4.2</u> : Let $x_1(0) = \beta_1$, $x_2(0) = \beta_2$. Then

$$\begin{vmatrix} u_1^0 \\ u_2^0 \end{vmatrix} = \begin{vmatrix} -1 \\ 0 \end{vmatrix} \qquad (\beta_1,\beta_2) \in \gamma_1^+$$

$$\begin{vmatrix} u_1^0 \\ u_2^0 \end{vmatrix} = \begin{vmatrix} 1 \\ 0 \end{vmatrix} \qquad (\beta_1,\beta_2) \in \gamma_1^-$$

$$\begin{vmatrix} u_1^0 \\ u_2^0 \end{vmatrix} = \begin{vmatrix} -1 \\ -1 \end{vmatrix} \qquad (\beta_1,\beta_2) \in \gamma_2^+ \tag{4.65}$$

$$\begin{vmatrix} u_1^0 \\ u_2^0 \end{vmatrix} = \begin{vmatrix} 1 \\ 1 \end{vmatrix} \qquad (\beta_1,\beta_2) \in \gamma_2^-$$

Let $\gamma_1 = \gamma_1^+ \cup \gamma_1^-$ $\quad \gamma_2 = \gamma_2^+ \cup \gamma_2^-$. The γ_1 and γ_2 curves are shown in Fig. 4.4. The regions R_1, R_2, R_3, R_4 are defined as follows.

$$R_1 = \left\{ (x_1,x_2) : -\frac{1}{2c_1} x_2^2 < x_1 < -\frac{1}{2(c_1+c_2)} x_2^2 \right\}$$

$$R_2 = \left\{ (x_1,x_2) : \frac{1}{2(c_1+c_2)} x_2^2 < x_1 < \frac{1}{2c_1} x_2^2 \right\} \tag{4.66}$$

$$R_3 = \left\{ (x_1,x_2) : x_1 > \beta_1 , (\beta_1,x_2) \in \gamma_2^- \ \gamma_1^+ \right\}$$

$$R_4 = \left\{ (x_1,x_2) : x_1 < \beta_1 , (\beta_1,x_2) \in \gamma_2^+ \ \gamma_1^- \right\} \tag{4.67}$$

<u>Lemma 4.3</u> : Only the following control strategy sequences are optimal to reach the origin:

$$\begin{vmatrix} 1 \\ 0 \end{vmatrix} , \begin{vmatrix} 1 \\ 1 \end{vmatrix} , \begin{vmatrix} -1 \\ 0 \end{vmatrix} , \begin{vmatrix} -1 \\ -1 \end{vmatrix} , \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} , \begin{vmatrix} -1 & -1 \\ 0 & -1 \end{vmatrix} , \begin{vmatrix} -1 & 1 \\ 0 & 0 \end{vmatrix} , \begin{vmatrix} 1 & -1 \\ 0 & 0 \end{vmatrix}$$

<u>Proof</u>: This we do by contradiction. Consider the sequence

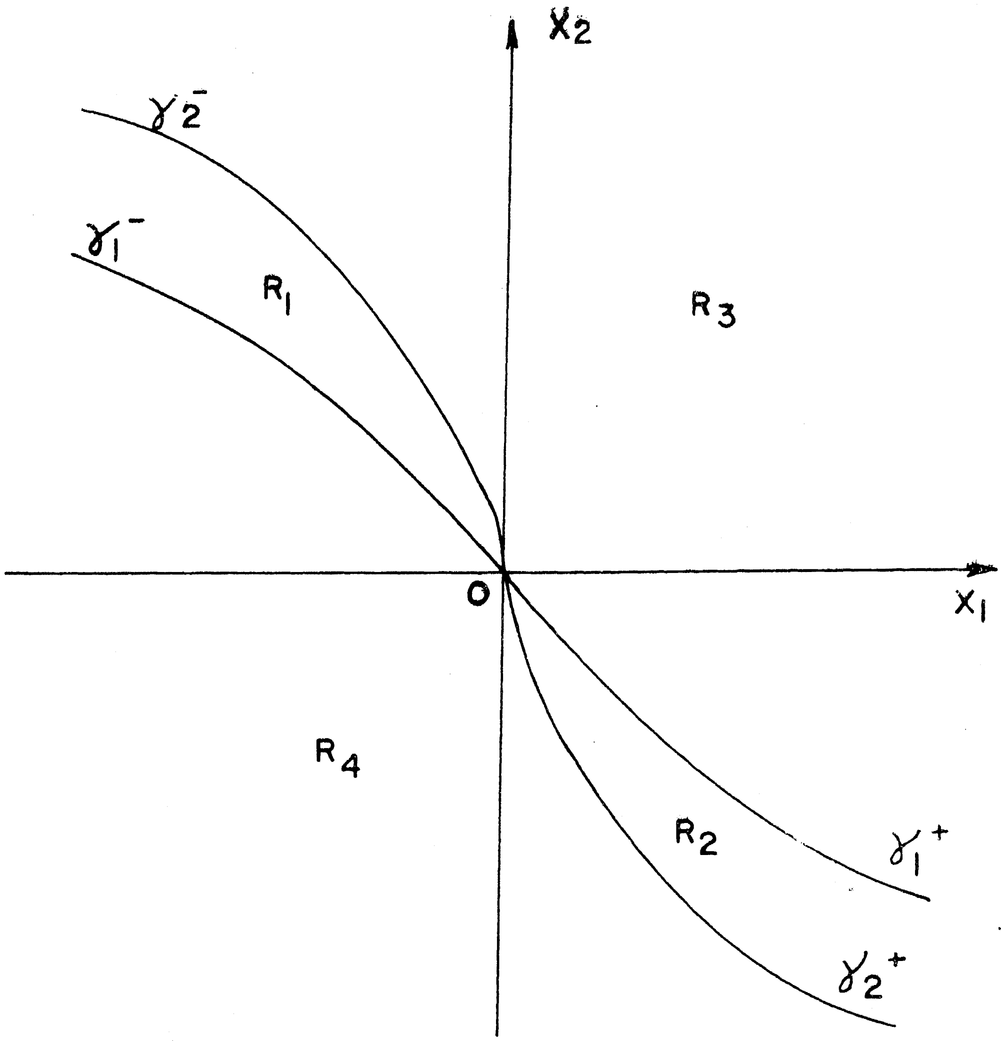


FIGURE 4.4 COMPLETE DECOMPOSITION OF THE $x_1 - x_2$ SPACE.

$\begin{vmatrix} -1 & -1 & 1 & 1 \\ -1 & 0 & 0 & 1 \end{vmatrix}$. The subsequence $\begin{vmatrix} -1 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix}$ is such that the corresponding trajector to reach the origin crosses γ_1^+ which is not allowed. Again with the subsequence $\begin{vmatrix} -1 & -1 & 1 \\ -1 & 0 & 0 \end{vmatrix}$, we start with $u_2 \neq 0$ and end up $|u_2| = 0$, at the origin. If $|u_2| = 0$ then Eq. (4.60 - 4.61) must be true for all times, which contradicts that $u_2 = 0$ at origin while $u_2 \neq 0$ at start.

Q.E.D.

Lemma 4.4 : The solution to the problem of determining optimal control strategies is

$$(\beta_1,\beta_2)\in R_1 \Rightarrow \begin{vmatrix} -1 & -1 \\ 0 & -1 \end{vmatrix}, \quad (\beta_1,\beta_2)\in R_2 \Rightarrow \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix}$$

$$(\beta_1,\beta_2)\in \gamma_2^- \Rightarrow \begin{vmatrix} -1 \\ -1 \end{vmatrix}, \quad (\beta_1,\beta_2)\in \gamma_2^+ \Rightarrow \begin{vmatrix} 1 \\ 1 \end{vmatrix}$$

$$(4.69)$$

$$(\beta_1,\beta_2)\in R_3 \Rightarrow \begin{vmatrix} -1 & 1 \\ 0 & 0 \end{vmatrix}, \quad (\beta_1,\beta_2)\in R_4 \Rightarrow \begin{vmatrix} 1 & -1 \\ 0 & 0 \end{vmatrix}$$

$$(\beta_1,\beta_2)\in \gamma_1^+ \Rightarrow \begin{vmatrix} 1 \\ 0 \end{vmatrix}, \quad (\beta_1,\beta_2)\in \gamma_1^- \Rightarrow \begin{vmatrix} -1 \\ 0 \end{vmatrix}$$

where (β_1,β_2) is the starting point in (x_1,x_2) space.

A similar analysis can be made with other criteria such as optimality relative to time and energy, or energy and fuel and the like. It is thus seen that in the extreme case with $c_1 = 0$, $c_2 \neq 0$, the results for fuel-optimal control will follow on the one hand and with $c_1 \neq 0$, $c_2 = 0$ the results of time-optimal control follow. Compare the corresponding cases in |51,82|.

Example 3 : N-Person Races : In this section we consider a preliminary investigation of ranking problems, where the rankings

are done at the end of a time interval, subject to differential and algebraic constraints. We envisage such ranking problems as differential games with participants. To the author's knowledge in literature hitherto there has been no discussion of ranking problems by differential games. Such ranking problems arise typically in races, athletic events and non-cooperative competition as in competitive public examinations, political campaigns, etc. Viewed as differential games, they are N-person differential games of kind; however, they can be converted into N-person differential games of degree [1, chapter 8]. As N-person differential games, both cooperative and non-cooperative versions can be considered. We restrict ourselves to the non-cooperative versions. In either case a second classification leads to the 'silent' race and 'noisy' race versions in differential games.

<u>The Silent Race</u> : The word 'silent' refers to the fact that each person has to control his position <u>unhindered</u> by other players and try to reach the set goal within his limitations.

<u>The Noisy Race</u> : If at least one player has his control over his position hindered by at least another player, then we refer to the race as 'noisy'.

In any track event, for example, the set goal for every player to start at $t = 0$ from a specified geographical point and to reach another specified geographical point. Let T^i refer to the time the ith player takes to reach the set goal. The 'umpire' then ranks the players as follows: let S_i be a subset of the players $\{1, 2, \ldots, N\} = \underline{N}$ such that

$$\operatorname*{Inf}_{i} \quad T^i \quad = \quad T_1 = T^k \quad k \in S_1,\ i \in \underline{N} \qquad (4.69)$$

Let $S_2 \subseteq (\underline{N} - S_1)$ such that

$$\operatorname*{Inf}_{i} \quad T^i \quad = \quad T_2 = T^k \quad k \in S_2,\ i \in (\underline{N} - S_1) \qquad (4.70)$$

and so on till the set $\underline{N}$ is exhausted. The ranking is then done as

$$S_1 \geq S_2 \geq S_3 \ldots \geq S_l \qquad l \leq N \qquad (4.71)$$

The obvious objective function for the ith player would then be to minimise T^i subject to the differential and algebraic constraints. These are now chosen. Since this is only a preliminary investigation, the differential equations and constraints chosen may not be realistic; they are only illustrative. We consider only second order dynamics for the players. [These may arise in the modelling of a runner as a system, perhaps, or the man-machine combination of a chauffer and a vehicle or the psychological and political factors in public examinations and elections. Other models could be tried as improved versions of the modelling here]. The game is then an N-person differential game under non-cooperation.

<u>The Silent Race Versions</u> : Given

<u>system dynamics</u>

$$\begin{aligned} \dot{x}_1^i &= x_2^i \\ \dot{x}_2^i &= A_2^i x_2^i + A_1^i x_1^i + B^i u^i + r^i \qquad i=1,\ldots,N \end{aligned} \qquad (4.72)$$

subject to the constraints

$$|x_2^i| \leq M^i \qquad M^i \in R \qquad (4.73)$$

$$0 \leq u^i \leq 1 \qquad i=1,\ldots,N \qquad (4.74)$$

where x_1^i = the linear position of the ith player.

x_2^i = the linear velocity of the ith player.

F^i = the initialisation force (supposed given).

u^i = the control action of player i .

The terminal goal (surface) :

$$x_1^i = L \qquad i=1,\dots,N \tag{4.75}$$

where L is the track length,

$$x_2^i = \text{free} \qquad T^i = \text{free} \tag{4.76}$$

Objective function :

$$T^{io} = \min_{|u^i|<1} \int_0^{T^i} dt \qquad i=1,\dots,N. \tag{4.77}$$

Solution : The Hamiltonian form of the N-Person differential game theory can be applied. The Hamiltonian is given by

$$H^i\left(x_1^i , x_2^i , u^i , \lambda_1^{ij} , \lambda_2^{ij} , i,j=1,\dots,N\right)$$

$$= \sum_{j=1}^{N} \lambda_1^{ij} x_2^j + \lambda_2^{ij} \left(A_2^j x_2^j + A_1^j x_1^j + B^j u^j + F^j\right) + 1 \tag{4.78}$$

The optimal control strategy for the ith player is obtained by minimising H^i in u^i, assuming $u^j = u^{jo}$, $j \neq i$, $j=1,\dots,N$, which implies

$$\begin{aligned} u^{io} &= 1 \qquad \text{if sign } \lambda_2^{ii} < 0 \\ &= 0 \qquad \text{otherwise.} \end{aligned} \tag{4.79}$$

The adjoint equations are given as

$$\frac{d\lambda_1^{ij}}{dt} = -A_1^j \lambda_2^{ij} \tag{4.80}$$

$$\frac{d\ \lambda^{ij}_2}{dt} = -\ \lambda^{ij}_1 - A^j_2\ \lambda^{ij}_2 - \text{sign}\ x_j \qquad (4.81)$$

with the boundary condition

$$\lambda^{ij}_2(T) = x^j_2(T) \ , \qquad \lambda^{ij}_1(T) = 0 \qquad (4.82)$$

<u>The Noisy Race Version</u> : Given

<u>system dynamics</u>

$$\begin{aligned} \dot{x}^i_1 &= x^i_2 \\ \dot{x}^i_2 &= A^i_1 x^i_1 + A^i_2 x^i_2 + \Sigma^N_{j=1} B^i_j u^i_j \ , \ i=1,\ldots,N. \end{aligned} \qquad (4.83)$$

where for $j \neq i$ B^i_j stands for the coefficient of interference in the dynamics of the ith player caused by the jth player and u^i_j the corresponding interference force with constraints

$$0 \leq u^i_j \leq 1 \qquad (4.84)$$

The constraints, terminal surface, payoffs remain as in (4.73), (4.75) to (4.77).

<u>Solution</u> : Again we can write the Hamiltonian as

$$H^i\left(x^i_1 \ , \ x^i_2 \ , \ u^i_j \ , \ \lambda^{ij}_1 \ , \ \lambda^{ij}_2 \ , \ i,j = 1,\ldots,N \right)$$

$$= \Sigma^N_{j=1} \left(\lambda^{ij}_1 x^j_2 + \lambda^{ij}_2 \ A^j_1 x^j_1 + A^j_2 x^j_2 + \Sigma^N_{k=1} B^j_k u^j_k \right) \qquad (4.85)$$

The strategies of player i are u^1_i , u^2_i , $\ldots, u^i_i, \ldots, u^N_i$. Since the objective of the ith player is to cause interference and trouble to other players, the only way he can achieve this is by requiring the Hamiltonian H^i to be a maximum in u^j_i

for $j \neq i$ and a minimum for $j = i$. This leads to

$$
\begin{aligned}
u_i^{jo} &= 0 \quad \text{if sign } B_i^j \lambda_2^{ij} < 0 \\
&= 1 \quad \text{otherwise}
\end{aligned}
\tag{4.86}
$$

for $i \neq j$, $i,j = 1,\ldots,N$, and

$$
\begin{aligned}
u_i^{io} &= 1 \quad \text{if sign } B_i^i \lambda_2^{ij} < 0 \\
&= 0 \quad \text{otherwise}
\end{aligned}
\tag{4.87}
$$

Thus he keeps track of the adjoint equations and interferes in the jth player's strategy as long as the corresponding optimal return gradient is positive while for his own dynamics the optimal force is applied when the corresponding optimal return gradient becomes negative. The adjoint differential equations are given as

$$
\begin{aligned}
\frac{d\lambda_1^{ij}}{dt} &= -A_1^j \lambda_2^{ij} \\
\frac{d\lambda_2^{ij}}{dt} &= -\lambda_1^{ij} - A_2^j \lambda_2^{ij} - \text{sign } x_j, \; i,j=1,\ldots,N
\end{aligned}
\tag{4.88}
$$

The boundary conditions are

$$
\lambda_2^{ij}(T) = x_2^j(T) , \quad \lambda_1^{ij}(T) = 0 , \; i,j=1,\ldots,N \tag{4.89}
$$

The umpire then applies the ranking procedure given in (4.69 - 4.71) to both the versions (the 'silent' and the 'noisy' cases).

These ranking problems can be considered similar to the games of timing, 'duels' etc. (The mathematical analysis of the 'cooperative' version of the races should be even more interesting as many sociological problems(such as bribing, threatening, cut throat competition, etc.) can be partially answered.)

4.4 CONCLUSIONS

This chapter has shown the possibilities afforded for the analysis of various system problems through N-person differential game theory. Though we have considered the non-cooperative case for the three problems, viz., the national economy model of two nations, the system design problem with two criteria and the 'races', the cooperative version should be even more interesting and worth investigating. This chapter also concludes our investigations of deterministic games. The remaining three chapters concern games with uncertainty in the game description, and many concepts investigated in deterministic games will be needed therein.

V GAMES WITH UNCERTAINTY

We now consider games where the parameter set W_q plays an important role. Each player is considered to have some lack of information about W_q and thus needs a correct assessment of W_q from instant to instant to implement the correct optimal policy. We term such games as Positional Games under Uncertainty. Such games arise in an engineering context in a variety of ways. A design problem for a computer controlled optimal system with uncertainty is one such instance. This is also an example of an off-line game. An on-line one-sided positional game with uncertainty is afforded by a human operator in a closed loop optimal control task (as in optimal maneuvering of planes). The human factors affect the two games differently. The engineering design of a man-machine complex to be built as a simulator for operational gaming provides another such example with on-line and off-line human factors playing an important role.

Such engineering situations with one or more human operators point to the need for a framework in games and decisions where both subjective and objective preferences can be handled. We use the model developed in chapter 2 for this framework. Decision making problems in systems engineering have been dealt by Hall and Machol |83, 84|; in the case of adaptive systems by Hsu and Meserve |85|, and Rosenfeld |86|; and in general in related areas by Ackoff |87|. Apart from decision making problems, other types of stochastic and adaptive positional games come under the purview of this chapter. A direct outcome of some of the conceptual

questions is provided by the Markov Positional game model in the next chapter. (See also Sakaguchi|33|,|58|).

5.2 GAMES AGAINST NATURE

One-sided games with only uncertainty present are termed games against nature. The positional game we now consider is $(X, U, Y, W_1, W_2, W_q, S, H, \Theta, I)$ where W_1 is the set of parameters of randomisations possibly chosen by the player, W_2 is the set of possibly known sets of parameters, either stochastic or distributed elements, W_q is the set of parameter about which very little is known. It is this parameter set which has an important bearing. The essence of the abstract problem is to specify the policy $u\ (t, w, y_s)$, $t \in \Theta$, $s \in [0, t)\ \Theta$, $w \in W_1 \times W_2$ such that a loss function specified in terms of x and u is minimized.

Various problems about the existence of such policies arise which somewhat parallel known problems in control theory. We list these briefly.

(i) What is the nature of the observability of the position and the unknown parameter $w_q \in W_q$, both off-line and on-line for the set Θ and as $\Theta \to R$?

(ii) What is the role of control policies on the observability of the process ?

(iii) Can the control policies be determined in terms of minimal information of position, parameters, and uncertainty ?

(iv) What shape do the policies take for unobservable parameters ?

(v) Is it possible to specify two policies such that one serves to minimise the loss function in terms of x and u_1, while the other minimises the loss function specified in terms of w_q and u_2, where u_1 and u_2 are the above two policies ?

Each of these questions needs to be examined in the presence of $w_1 \in W_1$, $w_2 \in W_2$. In many cases it turns out randomization is not required. So the set w_1 plays no role at all. The questions we have raised are not completely answered as we have to take the decision to restrict ourselves to answering these questions within the constraints of time and space.

Our next objective is to determine the set of dominant control policies. Let $u \in U \times W_1$, with the corresponding loss function as $I(\tilde{u})$. Let $\mathcal{I} = \{ I(\tilde{u}) : \tilde{u} \in U \times W_1, w_q \in W_q \}$ which can be partitioned into the subsets

$$\mathcal{I}_1 = \{ I(\tilde{u}) : \tilde{u} \in U \times W_1 \quad \text{for some fixed} \quad w_{q1} \in W_q \}$$
$$\mathcal{I}_2 = \{ I(\tilde{u}) : \tilde{u} \in U \times W_1 \quad \text{for some fixed} \quad w_{q2} \in W_q \}$$
$$\vdots$$
$$\mathcal{I}_p = \{ I(\tilde{u}) : \tilde{u} \in U \times W_1 \quad \text{for some fixed} \quad w_{qp} \in W_q \}$$

Since we start out with the assumption that $\mathcal{I}$ is bounded from below, $\mathcal{I}_1, \ldots, \mathcal{I}_p$ have an infimum say $I_1, I_2, \ldots, I_p$. Corresponding to each of these we have a candidate mixed strategy $\tilde{u}_1^*, \ldots, \tilde{u}_p^*$ none of which are preferable to each other since we do not know $w_{q1}, \ldots, w_{qp} \in W_q$ to determine the exact optimal strategy, we need an estimation or identification procedure. The partitions of the set $\mathcal{I}$ have been made into a finite number corresponding to each $w_q \in W_q$. This is done for convenience of discussion.

We have thus essentially determined optimal mixed strategies none of which dominate over each other only on the basis of the payoff function I. The given payoff function is unable to determine a unique optimal control strategy and this confronts us with a dilemma. To resolve this dilemma we need to examine in more detail about the properties of the set W_q and its relationship with the choice of optimal mixed control strategy.

In an off-line game, given enough computing facilities and time, we can sift through the set of all dominant strategies. The sifting procedures, however, depend on factors other than those specified. This is not possible in an on-line game where the dilemma ought to be resolved as the observations are made by a judicious use of the control actions and previous observations. Essentially, we need extra considerations based both on human factors introduced by the role of the decision maker and the physical system in the form of a super criterion. The statistical or stochastic averaging methods of a super criterion require a specified description, and these can be supplied by the extra procedures. [In Figs. 5.1 - 5.3, we depict in a block diagram description, the role of a human operator in on-line and off-line tasks.]

In literature |58-60| we find rational procedures for decision making under certainty, risk and uncertainty. In the case of the first two we require complete identification studies which can resolve this dilemma. These studies must then constitute off-line decision making. In the last case no such studies can be provided and, if any, they are insufficient. Hence we need additional procedures which are in the form of subjective

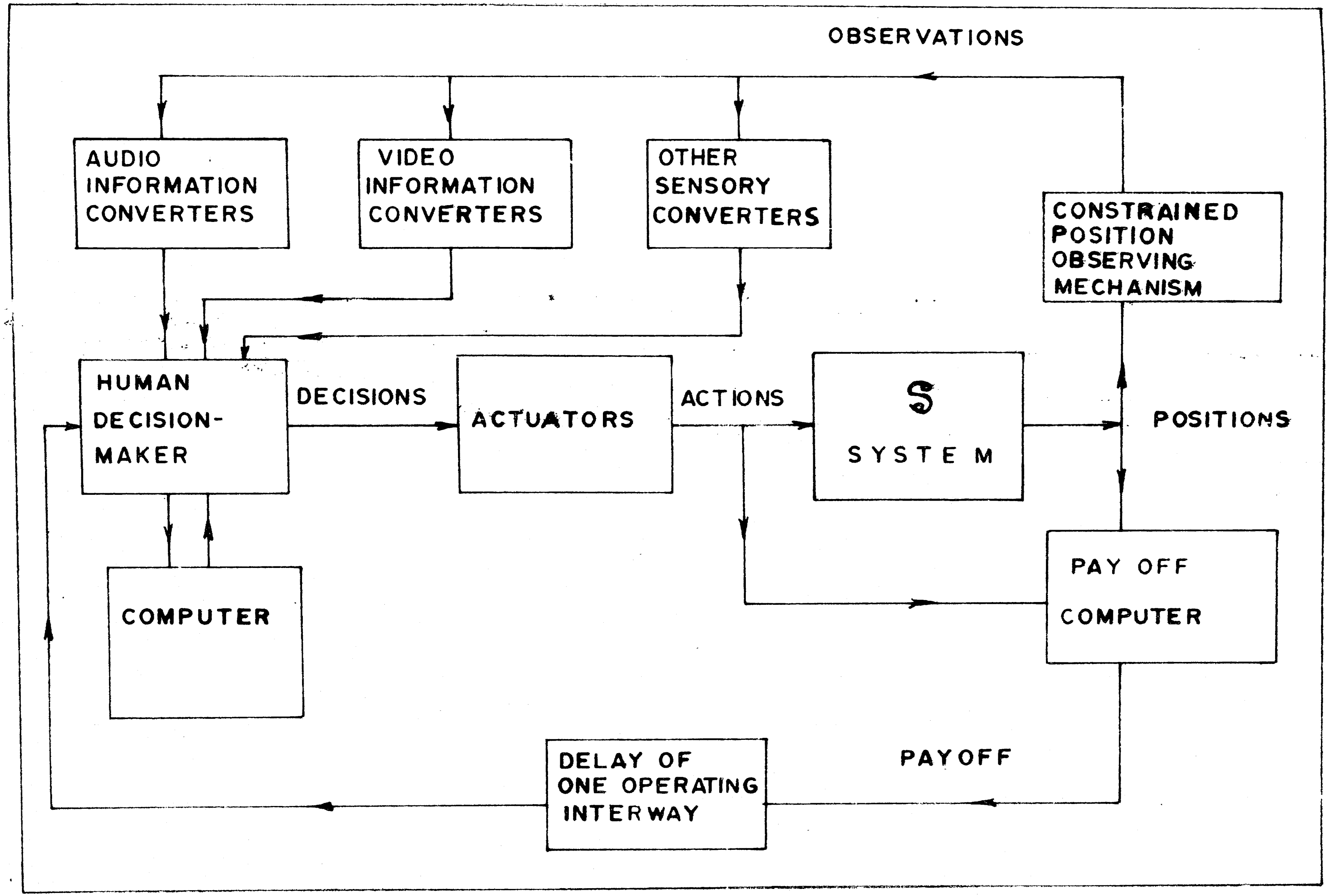


FIGURE 5·1 ON-LINE DECISION MAKING PROBLEM BY A HUMAN OPERATOR IN A CLOSED-LOOP TASK.

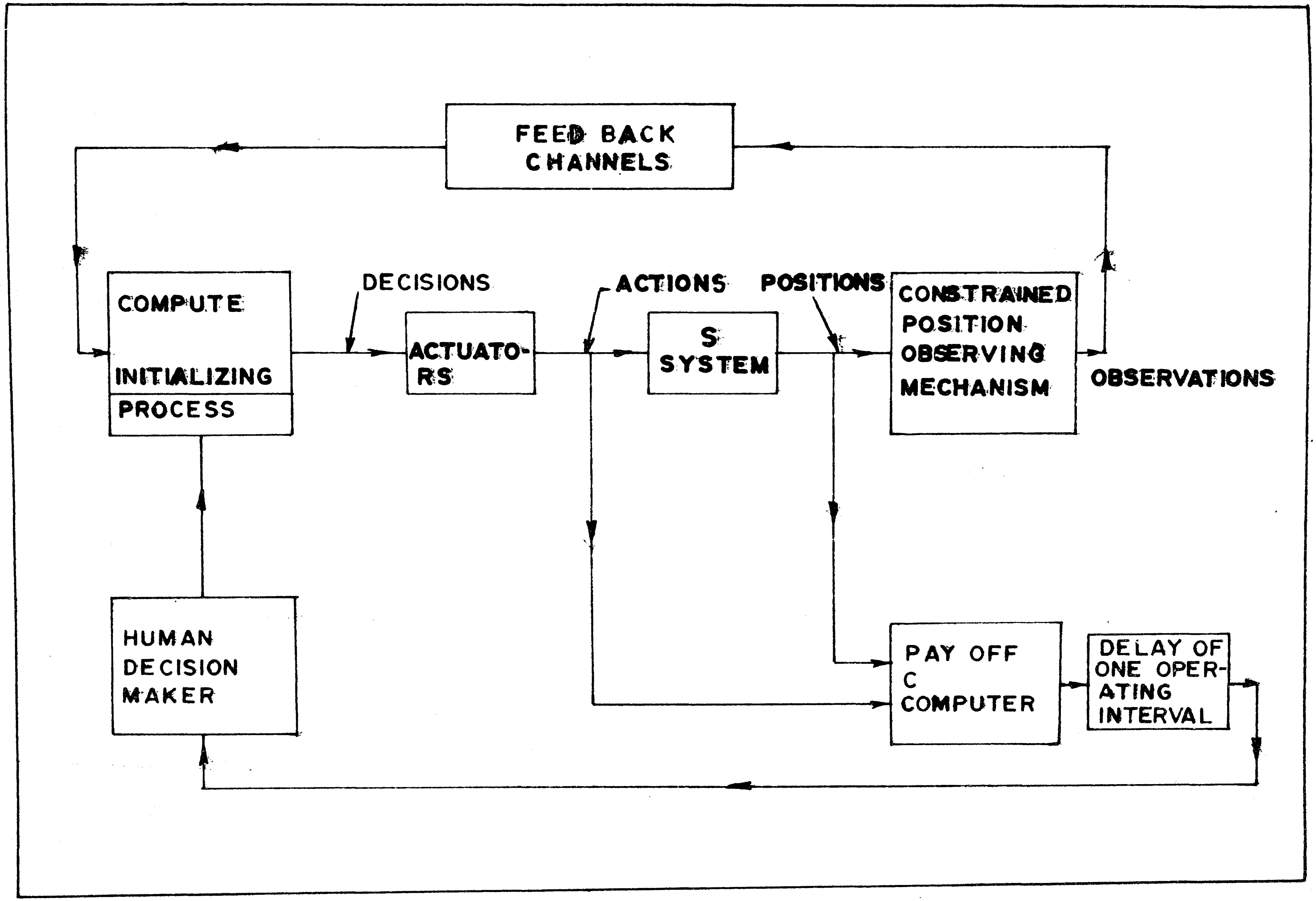


FIGURE 5·2 OFF-LINE DECISION MAKING PROBLEM BY A HUMAN DESIGNER.

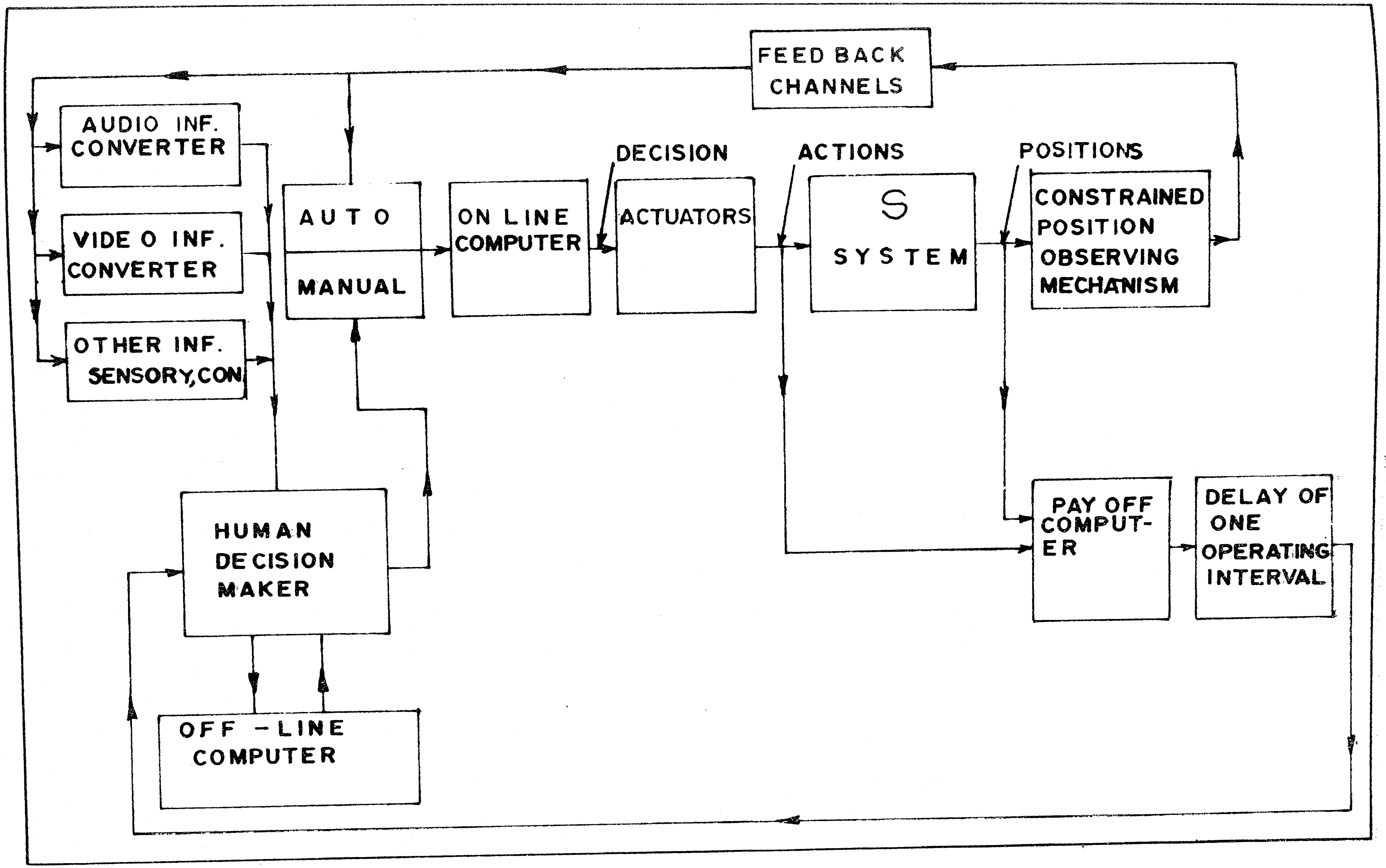


FIGURE 5.3 A MANUAL/AUTO DECISION MAKING BY A HUMAN-OPERATOR (MIXED MODE)

supercriteria. Subjective factors due to the presence of a human being in the loop (Fig. 8.1 - 8.3), would then require consideration of memory, recall, information, utility, etc. These will also arise if the designer has the problem posed to him in an ill-defined manner. One way is to reformulate the problem after making off-line studies. The other is to let his human factor come into play by selection of a priori assumptions and subjective factors.

Let us examine now the role of <u>supercriteria</u>|26|. Since a single ordering could not fix uniquely a mixed strategy u based on I, the decision-maker requires another criterion which he applies after I to the dominant set formed by the criterion function I. This is not specified in the given game. Perhaps, the criterion I and other objectives were rather in the nature of some platitudes which had no operational significance.

The decision-maker thus chooses the supercriteria entirely on a subjective level, i.e., dependent wholly on his likes and dislikes. Ideally a well-trained decision-maker picks the criterion based on experience. Thus the minimax or worst case optimal control policies are popular. Various measures of efficacies and efficiencies have been used(Sakaguchi |58|). In fact the second level criterion is so chosen as to definitely resolve the dilemma. Various measures of approximations have also been employed. Could we call these sub(-jective)-optimal ? Some classical supercriteria are listed in Appendix A. The conditions under which supercriteria should be used are listed in Appendix B.

Thus in an <u>adaptive-optimal control</u> problem where only an a priori knowledge has to be made available and a supercriterion

for resolving the dilemma, we certainly have to incur some loss when viewed retrospectively. This loss is precisely the extra 'cost' required to remove uncertainty. In this decision making problem under uncertainty, thus, the choice of the actual control law depends heavily on the current knowledge of the uncertainty parameter w_q and the supercriterion. The decision-maker (whether the designer or the human controller) has to make allowance for objective positions and subjective states to arrive at his decision. In this manner the control strategy depends on the subjective state vector trajectory, which represents the behaviour of the rate of removal of uncertainty in the decision-maker's mind. This state, by its very requirement, satisfies the Markov property which leads us next to enquire whether every supercriterion that resolves the dilemma does so by imbedding the system into a higher order Markov process. We believe this is so as is shown in the game considered in the next chapter. The various subjective factors and a priori assumptions merely serve to supply the initial conditions for this Markov process.

Let us also consider the relevance of subjectivity, objectivity and rationality |60-61|. When the designer (playing an off-line game) invokes supercriterion in the form of a priori probability distributions based on 'rational procedures', he is only indicating the degree of confidence he has in the nature of the parameter of uncertainty. As the designed system goes into operation the designer has provided some 'measure' built into the system such that the system automatically improves this degree of confidence. As against this, any objective factors have no need for improvement. These have essentially all past experience

contained in them. Thus when one refers to noisy measurements with known Gaussian noise, there is nothing one could do to improve this. The term rational players used in the theory of games is almost taken for granted ! Perhaps, in reality the consistent pattern of preferences of one player is not the same as that of other players|6|. It is in this connection that Harsanyi|60| considers rationality in games. The same is equally true in a dynamic control situation also.

Let us next consider problems with more than one decision-maker.

5.3 N(≥2)-PERSON GAMES

In a one-sided game against nature we had the following main factors: (i) certainty, (ii) risk (which were both off-line decision problems), (iii) uncertainty. To resolve the uncertainty a player resorted to subjective criteria. With two or more players and with uncertainty present, all the above factors are equally true for each player. Thus to one player the game may be with certainty or risk, to the other player it may be under uncertainty. The subjective criteria will also be different from those of the other players. The questions of controllability, observability and non-randomised policies have to be answered on the basis of different subjective and objective assumptions. In the context of positional games, therefore, much work remains to be done. One other source of uncertainty is the imperfect communication between partners of a player (as in Bridge, or between the commanding officer and his subordinates in armed forces.) We now describe some positional games with uncertainty.

(a) <u>The Trainer-Learner Problem</u> : In-flight decision making by a learner during an on-line task with the trainer beside him is a two-person game, cooperative to a certain extent. The flying of the plane to the trainer is mostly a routine task and hence from his point of view the modelling should be deterministic except for the unpredictable behaviour of the learner. To the learner, the process is new and hence the amount of information he gets at the end of the trial should be such as to maximise the total removal uncertainty. Hence the game appears to him as an adaptive-stochastic game under uncertainty. While the learner brings in certain amount of subjective factors, there are no such subjective factors in the trainer. The following game serves as a model for this problem.

Consider the positional game with incomplete structural information

$$\dot{x} = Ax + Bu + Cv \tag{5.1}$$

where the two players have different structural information sets. For the trainer (player I) controlling u, all the system coefficients are known, A, B, C and $x(t_0)$. For the learner (player II) all except the system matrix A are known. The common payoff function is to minimise in an interval $[0, T]$

$$I(u,v) = \langle x(T), Fx(T) \rangle + \int_0^T (\langle u(t), Qu(t) \rangle + \langle v(t), Rv(t) \rangle) \, dt \tag{5.2}$$

Player II knows A with an a priori subjective guess (probability) such that it is Gaussian with

$$E(A)|_0 \quad = \quad \bar{a}(0) \quad = \quad E(da(t))|_{t=0} \tag{5.3}$$

$$Var(A)|_0 \quad = \quad \Sigma(0) \quad = \quad Var(\ da(t)\ )|_{t=0} \tag{5.4}$$

The element of time enters into $\bar{a}$ and Σ due to the changing subjective probability densities. Thus the game appears to the two players differently. For player I (the trainer) the behaviour of the learner is only a realisation of the behaviour of a class of learners. He has, thus, a stochastic description for the behaviour through the flight period . This is given as

$$E(dv(t)) = \bar{v}(t)dt \tag{5.5}$$

$$Var(dv(t)) = S(t)dt \tag{5.6}$$

The system dynamics appears as

$$dx(t) \quad = \quad A(t)xdt + Bu(t)dt + Cdv(t) \tag{5.7}$$

To player II the incomplete structural information is now comprised of the a priori knowledge on A given by (5.3 - 5.4) and the known stochastic behaviour of the trainor

$$E(\ du(t)\ ) = \bar{u}(t)dt \tag{5.8}$$

$$Var(\ du(t)\ ) = P(t)dt \tag{5.9}$$

and the system dynamics appear as

$$dx(t) \quad = \quad da(t)\ x + Bd\bar{u}(t) + Cv(t)dt \tag{5.10}$$

He requires two criteria now - one is the objective function in (5.2) which he minimises by employing a subjective supercriterion

We thus see that it is possible to reduce certain games with incomplete information to games with multiple payoff functions to the players (some of which may be objective payoff functions and some subjective payoff functions), which may or may not have imperfect information but which have complete information, and the optimal control strategies are obtained by determining a Nash equilibrium point. Let us consider this formulation with a little more generality. Let the system constraints be

$$\dot{x} = G(x,u,v,w,t) \qquad x(t_o) = x_o \tag{5.15}$$

where x is the n-vector of positions

u is the r-vector of control actions of player I

v is the s-vector of control actions of player II

w is the vector of coefficients (parameters) characterising the uncertainty to player II. It could consist of structural coefficients.

Player I knows about w, while he does not know the control action of player II. Let the specified objective function of the game be

$$I(u,v) = \int_{t_o}^{T} f(x,u,v,w,t)dt \tag{5.16}$$

Further the synthesis of optimal strategies is required in terms of the observations

$$y = H_1(x,t) \tag{5.17}$$

$$z = H_2(x,w',t) \tag{5.18}$$

where w' is another parameter similarly not known to player II. Further player II has an average experience of the behaviour of player I. Thus, as the player II sees the game, he has to perform an estimation operation on-line, in order to determine (w and w'). It may turn out that from physical considerations (w and w') are constants but in the mind of player II these are variables (subjective variables) with a certain specified a priori probability distribution. He then invokes the following supercriterion to decide the estimation procedure

$$E(w \text{ and } w') = (w \text{ and } w')_{actual} \text{ at } t = T \quad (5.19)$$

$$\int_{t_0}^{T} \| \mathrm{Var}\,(w \text{ and } w') \| \, dt = \text{minimum} \quad (5.20)$$

Clearly this problem does not fit into the existing theory of differential games. In fact, we can mathematically consider the players as controlling separate systems with separate objective functions and overall constraints which are given below.

For player I, let the system dynamics be

$$\dot{x}^1 = G(x^1, u, v, \bar{v}, t) \quad (5.21)$$

$$y = H_1(x^1, t) \quad (5.22)$$

where $\bar{v}$ is known and v is unknown. An a priori specification on v is assumed

$$E(\dot{v}(t_0)) = \bar{v}(t_0) \quad (5.23)$$

with the objective function to extremise

$$\underset{u\in U}{\text{extremise}}\ I^1(u) = E_v\left\{\int_{t_0}^{T} f(x^1,u,v,\bar{w},t)dt\right\} \tag{5.24}$$

For player II, let the system dynamics be

$$\dot{x}^2 = G(x^2,u,v,w,t) \tag{5.25}$$

$$z = H_2(x^2, w',t) \tag{5.26}$$

where the unknowns to player II are w,w'. An a priori subjective probability distribution is assumed

$$E(w)\,|_{t=t_0} = \bar{W}(t_0) \tag{5.27}$$

$$E(w')|_{t=t_0} = \bar{W}'(t_0) \tag{5.28}$$

$$\text{and } Var(w)|_{t=t_0} = \bar{\bar{S}}(t_0) \tag{5.29}$$

$$Var(w')|_{t=t_0} = \bar{\bar{S}}(t_0) \tag{5.30}$$

Player II has two criteria - one is a modification of the given objective function by a supercriterion:

$$I_1^2(v) = E_{u,w,w'}\left\{\int_{t_0}^{T} f(x^2,u,v,w,t)dt\right\} \tag{5.31}$$

and the second is to minimise the integrated variance of the subjective random variables

$$I_2^2(v) = \int_{t_0}^{T} \|Var(w \text{ and } w')\|dt \tag{5.32}$$

The overall constraints are then specified as

$$\begin{aligned} x^1(t) &= x^2(t) \quad \text{for all } t\in[t_0,T] \\ x^2(t_0) &= x^1(t_0) = x_0 \end{aligned} \tag{5.33}$$

We can again split v into two components v^1 and v^2 and

consider these to be the control actions of the two agents of player II. Alternatively, these can be considered to be the control actions of two players. Thus the game can be reformulated as a complete information game for three players with specified a priori distributions for each of them:

$$\dot{x} = G(x,u,v^1,v^2,w,t) \tag{5.34}$$

$$y = H_1(x,t) \tag{5.35}$$

$$z_1 = H_2(x,w',t) = z_2 \tag{5.36}$$

The objective function of player I is now the same as in (5.24) with $x = x^1$, v replaced by v^1 and v^2. The objective function of the first agent of player II is $I_1^2(v^1)$ given by Eq. (5.31) with v^1 replacing v and x replacing x^2. The objective function of the second agent is given by Eq. (5.32) with v^2 replacing v. If an extension of Nash's non-cooperative theory to stochastic differential games is possible (and it seems feasible in view of chapter 4) then the above game can be solved within the framework of this extension.

Remarks: A question that arises is for what classes of games with incomplete structural information is this procedure valid ? Just as we could reduce certain games with incomplete position information to games with complete position information, can we convert any game with incomplete structural information to a game with complete information ? Perhaps of the various supercriteria, one which seems 'reasonable' is that of approximations. In this case the approximations can take the

form of a game with incomplete and imperfect information being approximated by a game with complete and perfect information. In such a procedure as pointed out by Isaacs |1|, we have to be wary of the advantages and dangers that accrue to a player from approximations. It may lead a player to believe that a strategy is more advantageous while its actual deployment leads to a loss for the player. Game theory as such enables us to infer the loss to a player when he plays non-optimally and the others play optimally. Nothing is said about what would happen when both players employ non-optimal strategies; not due to the player being not 'rational' but due to incomplete information.

We next consider a different type of game. Even though the game is of degree with random information, we can replace it by a game with complete perfect information which is a game of kind.

(b) Dynamical Modelling of a 'Dangerous' Game |62| : A dangerous game is defined as one in which threat expression may be unilateral, but punishment is always bilateral. In the former only one of the parties loses, in the latter both are losers substantially. Such games prevail in society: a headlong crisis in marital affairs of a couple, the threat of thermo-nuclear war between the two giant nations, two chauffers under the influence of alcohol about to crash on a highway crossing. Psychologists have tried to build simulators where the structurally similar properties are studied. One such simulator built with remote control equipment has been described by Swingle |62|. It consists of two toy trains A and B on a

parallel circular track. Refer to Fig. 5.4. The trains are started from the point marked START. Region C is a 'tunnel' area. The rules of the game are simple. The players watch a 'go' signal which is given at random times. Then, the objective of each player is to get into and out of the tunnel first before the other player enters and the player that comes out first unscathed wins a point from the other. If both are in the tunnel simultaneously, then both lose all their accumulated points. The winner is the person with the largest number of points, after a fixed number of runs not specified to the players. The constraints on the game are the speeds of the trains which are restricted to a maximum of V_m inches/min. We can identify two games here - one a differential game of kind for each run and the other a positional difference game of degree for the overall game. We shall formulate the two. A block diagram of the game is given in Fig. 5.5.

(i) <u>A Differential Game of Kind</u>: Refer to the following idealisation in Fig. 5.6. We consider only a circular track for the trains. The tunnel is the sector marked C. Let θ_1 denote the angular position of player I and θ_2 the angular position of player II from O. The dynamics of the two trains are given as

$$\begin{aligned} \ddot{\theta}_1 + a_1\dot{\theta}_1 &= u_1 \qquad \theta_1(0) = \dot{\theta}_1(0) = 0 \\ \ddot{\theta}_2 + a_2\dot{\theta}_2 &= u_2 \qquad \theta_2(0) = \dot{\theta}_2(0) = 0 \end{aligned} \qquad (5.37)$$

u_1 and u_2 are the controlling forces of the two players constrained by

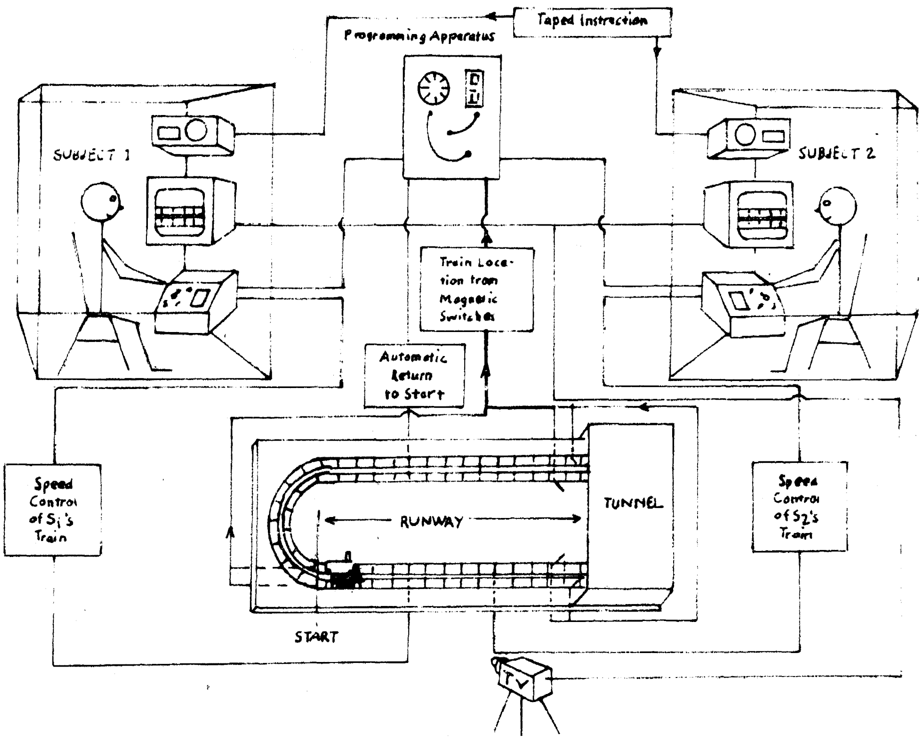


FIGURE 5·4 SWINGLE'S SET UP FOR THE 'DANGEROUS GAME'

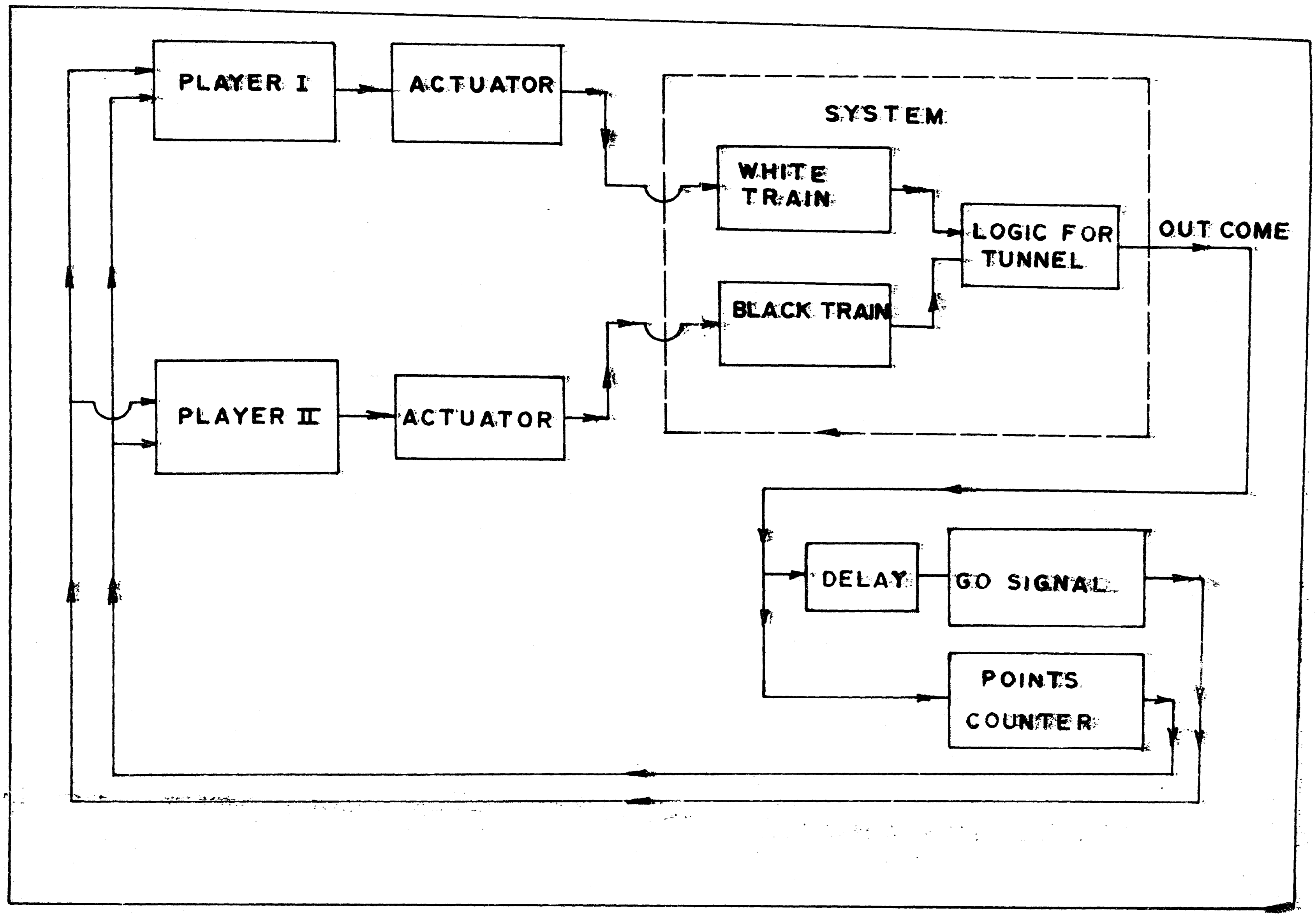


FIGURE 55 A BLOCK DIAGRAM FOR THE CONFLICT SIMULATOR.

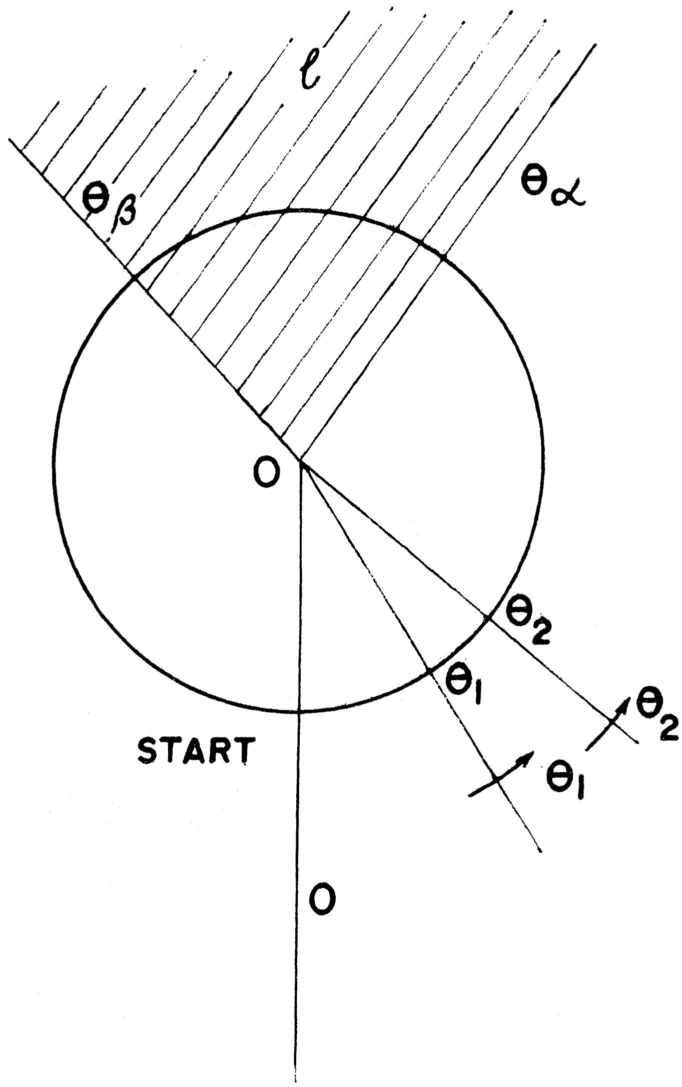


FIGURE 5.6 MATHEMATICAL IDEALIZATION OF THE SITUATION IN FIGURE 5.4.-5.5.

$$|u_1| \leq 1 \qquad |u_2| \leq 1 \tag{5.38}$$

$$\begin{aligned} &u_1(\tau_1) = 0 \qquad 0 < \tau_1 \quad \tau_1 \text{ being random} \\ &u_2(\tau_2) = 0 \qquad 0 < \tau_2 \quad \tau_2 \text{ being random} \end{aligned} \tag{5.39}$$

The payoff to player I is

+1 if he gets out of C before player II gets in

−1 if he gets into C after player II gets out of it

0 if both are in C simultaneously.

The problem also has bounded state (position) variables:

$$|\dot{\theta}_1| \leq \gamma_1 \qquad |\dot{\theta}_2| \leq \gamma_2 \tag{5.40}$$

Let us consider $\tau_1 < \tau_2$ then the problem can be considered equivalently of one with random initial conditions for player I starting at instant τ_2. $\theta_1(0)$ is then a random variable $\theta_1(0) \in [0, \theta_2]$, $\dot{\theta}_1(0)$ is a random variable $\dot{\theta}_1(0) \in [0, \gamma_1]$. A simplifying approximation could be made by considering a deterministic differential game of kind, with one of the players always having a lead in the angular position and velocity. There is no element of uncertainty here. Suppose instead we consider that one of the players has in addition a random input as, say,

$$\ddot{\theta}_1 + a_1\dot{\theta}_1 = u_1 + \xi_1 \tag{5.41}$$

or he has random observations

$$y_1 = \theta_2 + \eta \tag{5.42}$$

what are the methods of solving the stochastic differential games of kind ?

(2) A Positional Difference Game of Degree : We now consider the above game when many runs are conducted. Let state 1 denote that player I has crossed the tunnel first unscathed. Let state 0 refer to the collided state. Let state 2 denote the situation when the second player escapes unscathed. If player I finds himself at state 1 at the n-th run, he has now a choice to make a transition to 1, 0 or 2. However, these transitions are governed by a conditional probability transition matrix whose elements the players control. Let $v_i(n)$ denote the cumulative points player I receives when he finds himself at the ith state at the instant n. Then, we can write the cumulative points growth vector for the player I as

$$\begin{bmatrix} v_1(n+1) \\ v_0(n+1) \\ v_2(n+1) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + \begin{bmatrix} r_{11} & r_{10} & r_{12} \\ r_{01} & r_{00} & r_{02} \\ r_{21} & r_{20} & r_{22} \end{bmatrix} \begin{bmatrix} v_1(n) \\ v_0(n) \\ v_2(n) \end{bmatrix} \tag{5.43}$$

where $R = [r_{ij}]$ is a matrix which is state dependent. If player I is in state 1 at instant n then

$$R = R_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \tag{5.44}$$

If he is in state 0 at n, $R = R_0 = \underline{0}$

If he is in state 2 at n,

$$R = R_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5.45}$$

The psychology of the players demands neither one to accumulate a large number of points. Hence if $|v_1(n)|$ is large the transition must be heavily weighed towards state 0. If they are at state 0 at n, it hardly matters whether at $n+1$, the state is 1, 0 or 2. Hence the transition probabilities are equal, $p_{10} = 1/3$. The following transition matrix reflects the considerations of each player's behaviour:

$$P(n) = \begin{bmatrix} ye^{-|v_1(n)|} & 1/3 & ye^{-|v_2(n)|} \\ 1-(y+z)e^{-|v_1(n)|} & 1/3 & 1-(y+z)e^{-|v_2(n)|} \\ ze^{-|v_1(n)|} & 1/3 & ze^{-|v_2(n)|} \end{bmatrix} \qquad (5.46)$$

The strategy variable $y(n)$ is controlled by player I and the strategy variable z by player II and these are constrained as

$$\begin{aligned} 0 &\leq y(n) \leq y_m \\ 0 &\leq z(n) \leq z_m \end{aligned} \qquad (5.47)$$

$$\begin{aligned} ye^{-|v_1(n)|} &< 1 \\ ze^{-|v_1(n)|} &< 1 \end{aligned} \qquad (5.48)$$

Thus we have time varying, state dependent dynamics. The strategy variables are also time variable and state constrained.

Let n_s^i denote the number of runs actually required to reach the state 0 first starting at state i. Then, analogous to differential games, we pose the following three problems:

(i) Both parties cooperate:

Find strategies $y(n)$, $z(n)$ such that the constraints (5.43 - 5.48) are satisfied and such that

$$(n_s^i)^* \quad = \quad \max_{y(n)} \; \max_{s(n)} \; n_s^i \tag{5.49}$$

(ii) One person cooperates and the other does not. Then, subject to the same constraints, find

$$\overline{n_s^i} \quad = \quad \min_{y(n)} \; \max_{s(n)} \; n_s^i \tag{5.50}$$

(iii) Both behave dangerously. Then, subject to the same constraints, find

$$(n_s^i)_* \quad = \quad \min_{y(n)} \; \min_{s(n)} \; n_s^i \tag{5.51}$$

This game can be solved by making use of the theory of Markov chains. We only give here, however, broad features of the solution.

In case (i) the best strategy at any instant for player I is to use

$$\inf \left\{ e^{|v_1(n)|},\, y_m \right\} \tag{5.52}$$

and for player II is to use

$$\inf \left\{ e^{|v_1(n)|},\, s_m \right\} \tag{5.53}$$

In case (ii), if the first player wants to collide at any instant he should choose the strategy

$$y \quad = \quad 0 \tag{5.54}$$

while the other player continues to choose the strategy

$$\inf \left\{ e^{|v_1(n)|},\, s_m \right\} \tag{5.55}$$

In case (iii), both the players choose the strategy

$$y = 0 \; , \; z = 0 \tag{5.56}$$

We next determine the maximum expected reward a player gets in all the three cases. Consider player I. His maximum expected reward is given by summing over the rewards multiplied by the probability that it could be the reward. Thus starting from state 1 the maximum expected reward is given by

$$
\begin{aligned}
R = R = \; & n_s \times \left[\prod_{n=1}^{n_s} y_n e^{-|v_1(n)|} \right] \\
& + n_{s-1} \times \left[\prod_{n=1}^{n_{s-1}} y_n e^{-|v_1(n)|} \; z_n e^{-|v_1(n)|} \right. \quad + \\
& \qquad \sum_{j=1}^{n_{s-1}} \prod_{n=1}^{j} y_n \, z_{j+1} \prod_{n=j+2}^{n_s} y_n e^{-|v_1(n)|} \\
& \qquad \left. + \text{ similar terms} \right] + \text{similar expressions.}
\end{aligned}
\tag{5.57}
$$

where we substitute for n_s appropriately $(n_s^i)^*$, $(n_s^i)_*$, $\bar{n}_s^i$.

We conclude our investigation of some real life problems here. The next section deals with explicit methods to determine optimal strategies for linear stochastic positional games under different conditions of incomplete information. This will also provide a link to the problem considered in chapter 3, section 5.

5.4 LINEAR STOCHASTIC POSITIONAL GAMES

In this section we shall determine explicit control strategies for the players and a partial-differential equation for the Value function for positional games with complete position information. This will serve as a guide line for solving positional games under uncertainty once these are reduced to an equivalent stochastic differential game or positional game. In the next section we consider the positional game with random partial information.

(a) Complete Information : We are given that the two players control the position vector of a system through the differential equation

$$dx(t) = (A(t)dt + d\alpha(t))x(t) + (B(t)dt + d\beta(t))u(t) + (C(t)dt + d\gamma(t))v(t) + r(t)dt + d\eta(t) \tag{5.88}$$

$$x(t_0) = x_0$$

where $x(t)$ is the n-vector of positions at time t

$u(t)$ is the r-vector of control actions of player I at time t

$v(t)$ is the s-vector of control actions of player II at time t

$(A(t)dt + d\alpha(t))$ is the random system matrix of order $n \times n$

$(B(t)dt + d\beta(t))$ is the random gain matrix of order $n \times r$ for player I

$(C(t)dt + d\gamma(t))$ is the random gain matrix of order $n \times s$ for player II.

The random elements are assumed to be Wiener processes with independent increments whose mean and variance processes are given a priori as

$$\begin{aligned} &E(\, d\alpha(t)\,) = 0 \qquad && E(\, d\beta(t)\,) = 0 \\ &E(\, d\eta(t)\,) = 0 \qquad && E(\, d\gamma(t)\,) = 0 \end{aligned} \tag{5.59}$$

and

$$\begin{aligned} &E(\, (d\alpha(t))_i\, (\, d\alpha(t)\,)_j^T\,) = A_{ij}(t)dt \qquad i,j=1,\dots,n \\ &E(\, (d\beta(t))_i\, (\, d\beta(t)\,)_j^T\,) = B'_{ij}(t)dt \qquad i,j=1,\dots,r \\ &E(\, (d\gamma(t))_i\, (\, d\gamma(t)\,)_j^T\,) = C'_{ij}(t)dt \qquad i,j=1,\dots,s \\ &E(\, d\eta(t)\, (\, d\eta(t)\,)^T\,) = S(t)dt \end{aligned} \tag{5.60}$$

where $(\ )_i$ denotes the ith column of $(\ )$. The payoff function to player II from player I is given by

$$I(u,v) = E\left\{ \int_{t_0}^{T} (\langle u(t),Ru(t)\rangle - \langle v(t),Sv(t)\rangle)dt + \langle x(T),Fx(T)\rangle \right\} \tag{5.61}$$

We have to determine optimal control strategies of the two players such that $I(u,v)$ has a saddle-point in pure strategies:

$$I(u^0,v) \leq I(u^0,v^0) \leq I(u,v^0) \tag{5.62}$$

where u and v are chosen from the constraint sets

$$U = \left\{ u_i(t):\ |u_i(t)| \leq 1 \ \text{ for each } t\in[t_0,T],\ i=1,\dots,r \right\} \tag{5.63}$$

$$V = \left\{ v_i(t):\ |v_i(t)| \leq 1 \ \text{ for each } t\in[t_0,T],\ i=1,\dots,s \right\} \tag{5.64}$$

Let us compute the quantities $E\big(dx(t)\big)$ and $E\big(dx(t)dx^T(t)\big)$

which we shall need:

$$E(dx(t)) = (A(t)E(x(t)) + B(t)u(t) + C(t)v(t) + r(t))dt + O(dt) \tag{5.65}$$

$$\begin{aligned} E(dx(t)dx^T(t)) = \; & E(\,d\alpha(t)x(t)x^T(t)d\alpha^T(t)\,) + \\ & E(\,d\beta(t)u(t)u^T(t)d\beta^T(t)\,) + \\ & E(\,d\gamma(t)v(t)v^T(t)d\gamma^T(t)\,) + \\ & E(\,d\eta(t)d\eta^T(t)\,) + O(dt) \end{aligned} \tag{5.66}$$

We now proceed to determine the optimal strategies by dynamic programming techniques. Let

$$\begin{aligned} W(x_o,t_o) &= \min_{u(t)\in U}\ \max_{v(t)\in V}\ I(u,v) = \max_{v(t)\in V}\ \min_{u(t)\in U}\ I(u,v) \\ &= \min_{u(t)\in U}\ \max_{v(t)\in V}\ E\Big\{\langle x(T),Px(T)\rangle + \\ &\qquad \int_{t_o}^{T} (\langle u(t),Ru(t)\rangle - \langle v(t),Sv(t)\rangle)dt\Big\} \end{aligned} \tag{5.67}$$

By the principle of imbedding and the principle of optimality we can write

$$\begin{aligned} W(x_o,t_o) = \min_{u(t_o)\in U}\ \max_{v(t_o)\in V}\ & E\Big\{(\langle u(t),Ru(t)\rangle - \langle v(t),Sv(t)\rangle)dt \\ & + W(x_o+dx_o,t_o+dt_o)\Big\} \end{aligned} \tag{5.68}$$

with $\quad W(x(T),T) = \langle x(T),Px(T)\rangle \qquad (5.69)$

On expanding $W(x_o + dx_o, t_o + dt_o)$ by Taylor's series and retaining only the first order terms in dt we have after rearranging

$$\frac{-\partial W(x_0,t_0)}{\partial t_0}$$

$$= \min_{u(t_0)\in U} \max_{v(t_0)\in V} E\Big\{\langle u(t_0), Ru(t_0)\rangle - \langle v(t_0), Sv(t_0)\rangle$$

$$+ \Big\langle \frac{\partial W(x_0,t_0)}{\partial x_0}, A(t)x(t) + B(t)u(t) + C(t)v(t) + r(t) \Big\rangle$$

$$+ \frac{1}{2}\frac{d}{dt}\Big\{ tr\Big\{ \frac{\partial^2 W(x_0,t_0)}{\partial x_0^2} \big(d\alpha(t)x(t)x^T(t)d\alpha^T(t)$$

$$+ d\beta(t)u(t)u^T(t)d\beta^T(t)$$

$$+ d\gamma(t)v(t)v^T(t)d\gamma^T(t)$$

$$+ d\eta(t)d\eta^T(t) \big)\Big\}\Big\} + O(dt) \Big\} \quad (5.70)$$

where $\frac{\partial^2 W(x_0,t_0)}{\partial x_0^2}$ stands for the Hessian matrix (second partial matrix) of $W(x_0,t_0)$ with respect to x_0. Let us identify $x_0 = x$, $t_0 = t$ and further define

$$\frac{d}{dt} E\Big(\frac{1}{2} tr\Big(\frac{\partial^2 W(x,t)}{\partial x^2} \; d\alpha(t)x(t)x^T(t)d\alpha^T(t) \Big)\Big) = \langle x(t), \underline{g}(t)x(t)\rangle \quad (5.71)$$

$$\frac{d}{dt} E\Big(\frac{1}{2} tr\Big(\frac{\partial^2 W(x,t)}{\partial x^2} \; d\beta(t)u(t)u^T(t)d\beta^T(t) \Big)\Big) = \langle u(t), R_1(t)u(t)\rangle \quad (5.72)$$

$$\frac{d}{dt} E\Big(\frac{1}{2} tr\Big(\frac{\partial^2 W(x,t)}{\partial x^2} \; d\gamma(t)v(t)v^T(t)d\gamma^T(t) \Big)\Big) = \langle v(t), S_1(t)v(t)\rangle \quad (5.73)$$

Then Eq. (5.70) can be rewritten

$$-\frac{\partial W(x,t)}{\partial t} = \min_{u(t)\in U} \max_{v(t)\in V} \{ \langle u(t), (R+R_1(t)) u(t)\rangle + \langle x(t), Q(t) x(t)\rangle$$

$$-\langle v(t), (S-S_1(t)) v(t)\rangle + \frac{1}{2} \operatorname{tr} \{ \frac{\partial^2 W(x,t)}{\partial x^2}$$

$$+\langle \frac{\partial W(x,t)}{\partial x}, A(t)x(t)+B(t)u(t)+C(t)v(t)+r$$

(5.74)

Let us consider for the moment that u, v are unconstrained, $U = \mathcal{R}^r$, $V = \mathcal{R}^s$. Then on differentiating (5.74) with respect to u and v and on rearranging we have at the optimal point

$$(R+R_1(t)) u(t) = -B^T(t) \frac{\partial W(x,t)}{\partial x} \tag{5.75}$$

$$(S-S_1(t)) v(t) = C^T(t) \frac{\partial W(x,t)}{\partial x} \tag{5.76}$$

or

$$u^o = -(R + R_1(t))^{-1} B^T(t) \frac{\partial W(x,t)}{\partial x} \tag{5.77}$$

$$v^o = (S - S_1(t))^{-1} C^T(t) \frac{\partial W(x,t)}{\partial x} \tag{5.78}$$

In order that we are assured of a saddlepoint in Eq. (5.62) we now place restrictions on $(R + R_1(t))$ and $(S-S_1(t))$. We require that both these matrices be positive definite for each $t\in[t_0,T]$. If we now consider the constraints given in (5.63 – 5.64) we can write

$$u^o = -SAT ((R+R_1(t))^{-1} B^T(t) \frac{\partial W(x,t)}{\partial x}) \tag{5.79}$$

$$v^o = SAT ((S-S_1(t))^{-1} C^T(t) \frac{\partial W(x,t)}{\partial x}) \tag{5.80}$$

where the function SAT is used in the same sense as in chapter 3, i.e., it is the vector saturation function. In the unconstrained

case on substituting Eqs. (5.77 - 5.78) into Eq. (5.74) we have

$$
\begin{aligned}
\frac{-\partial W(x,t)}{\partial t} = & -\frac{1}{2} < \frac{\partial W(x,t)}{\partial x}, B(t)\ (R+R_1(t))^{-1}\ B^T(t)\ \frac{\partial W(x,t)}{\partial x} > \\
& + \frac{1}{2} < \frac{\partial W(x,t)}{\partial x}, C(t)\ (S-S_1(t))^{-1} C^T(t) \quad \frac{\partial W(x,t)}{\partial x} > \\
& + <x(t), \underline{g}(t)x(t)> + \frac{1}{2}\ tr\left\{ \frac{\partial^2 W(x,t)}{\partial x^2}\ \mathcal{E} \right\} \\
& + < \frac{\partial W(x,t)}{\partial x}, A(t)x(t) + r(t) >
\end{aligned}
\tag{5.81}
$$

To solve Eq. (5.81) we assume that $W(x,t)$ has the separable form

$$
W(x,t) = k_0(t) + <\mathbf{K}(t), x> + < x, K(t)x > \tag{5.82}
$$

Then

$$
\frac{\partial W(x,t)}{\partial x} = \mathbf{K}(t) + 2K(t)x \tag{5.83}
$$

$$
\frac{\partial^2 W(x,t)}{\partial x^2} = 2K(t) \tag{5.84}
$$

Hence by substituting (5.82 - 5.84) into (5.81) we have

$$
\begin{aligned}
-\frac{dk_0(t)}{dt} = & \ tr\{ K(t)\ \mathcal{E} \} + < \mathbf{K}(t), r(t) > \\
& - \frac{1}{2} < \mathbf{K}(t), (B(t)(R+R_1(t))^{-1}\ B^T(t) \\
& \qquad - C(t)(S-S_1(t))^{-1}\ C^T(t))\ \mathbf{K}(t) >
\end{aligned}
\tag{5.85}
$$

$$
\begin{aligned}
-\frac{d\mathbf{K}(t)}{dt} = & \ (\ A(t)+C(t)(S-S_1(t))^{-1}\ C^T(t)-B(t)(R+R_1(t))^{-1}\ B^T(t)\ )^T \mathbf{K}(\\
& + 2K(t)\ r(t)
\end{aligned}
\tag{5.86}
$$

$$-\frac{dK(t)}{dt} = K(t)\,A(t) + A^T(t)\,K(t) + \underline{g}(t) - K(t)\left(B(t)(R+R_1(t))^{-1}\,B^T(t) - C(t)(S-S_1(t))^{-1}C^T(t)\right)K(t) \tag{5.87}$$

with the boundary conditions

$$k_0(T) = 0\,, \qquad \underline{k}(T) = 0\,, \qquad K(T) = F \tag{5.88}$$

Substituting (5.85) into (5.77 - 5.78) we get

$$u^0 = -(\,R + R_1(t)\,)^{-1}\,B^T(t)\;(\,\underline{k}(t) + 2K(t)x\,) \tag{5.89}$$

$$v^0 = (\,S - S_1(t)\,)^{-1}\,C^T(t)\;(\,\underline{k}(t) + 2K(t)x\,) \tag{5.90}$$

In the case of constraints, a corresponding partial differential equation has to be solved to obtain the Value function. We exemplify the complete information stochastic game through the following example. Let the position be governed by

$$\begin{bmatrix} dx_1(t) \\ dx_2(t) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} dt + \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} d\beta_1 \\ d\beta_2 \end{bmatrix} \right\} u(t) + \left\{ \begin{bmatrix} -1 \\ -2 \end{bmatrix} + \begin{bmatrix} d\gamma_1 \\ d\gamma_2 \end{bmatrix} \right\} v(t) + \begin{bmatrix} d\eta_1(t) \\ d\eta_2(t) \end{bmatrix} \tag{5.91}$$

$$x(0) = x_0 \text{ specified.}$$

Let

$$E \begin{bmatrix} d\beta_1(t) \\ d\beta_2(t) \end{bmatrix} \begin{bmatrix} d\beta_1(t) & d\beta_2(t) \end{bmatrix} = \begin{bmatrix} b_1 & 0 \\ 0 & b_2 \end{bmatrix} dt \tag{5.92}$$

$$E \begin{bmatrix} d\gamma_1(t) \\ d\gamma_2(t) \end{bmatrix} \begin{bmatrix} d\gamma_1(t) & d\gamma_2(t) \end{bmatrix} = \begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix} dt \tag{5.93}$$

$$E\begin{bmatrix} d\eta_1(t) \\ d\eta_2(t) \end{bmatrix}\begin{bmatrix} d\eta_1(t) & d\eta_2(t) \end{bmatrix} = \begin{bmatrix} h_1 & 0 \\ 0 & h_2 \end{bmatrix} dt \qquad (5.94)$$

Determine saddlepoint strategies for the following payoff function subject to (5.91 - 5.94).

$$I(u,v) = E\left(\int_0^T (ru^2 - sv^2)dt + f_1x_1^2(T) + f_2x_2^2(T) \right) \qquad (5.95)$$

Let
$$W(x,t) = \begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} k_{11}(t) & k_{12}(t) \\ k_{12}(t) & k_{22}(t) \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$+ z_1(t)\, x_1 + z_2(t)\, x_2 + k_0(t) \qquad (5.96)$$

Now

$$\frac{1}{2}\,\mathrm{tr}\left\{ \begin{bmatrix} k_{11}(t) & k_{12}(t) \\ k_{12}(t) & k_{22}(t) \end{bmatrix}\begin{bmatrix} b_1 & 0 \\ 0 & b_2 \end{bmatrix} u^2 \right\} \begin{matrix} d \\ dt \end{matrix} = r_1(t)u^2dt \qquad (5.97)$$

Hence

$$r_1(t) = b_1k_{11}(t) + b_2k_{22}(t) \qquad (5.98)$$

Similarly

$$s_1(t) = c_1k_{11}(t) + c_2k_{22}(t) \qquad (5.99)$$

Therefore, the optimal control strategies are written:

$$u^0(t) = -(r+b_1k_{11}(t)+b_2k_{22}(t))^{-1}\,(z_1(t)+z_2(t)+(k_{11}(t)+k_{12}(t))x_1$$
$$+(k_{12}(t)+k_{22}(t))x_2)$$
$$(5.100)$$

$$v^0(t) = (s-c_1k_{11}(t) - c_2k_{22}(t))^{-1}\,(-z_1(t) - 2z_2(t)$$
$$-(k_{11}(t)+2k_{12}(t))x_1 - (k_{12}(t)+2k_{22}(t))x_2)$$
$$(5.101)$$

where

$$-\dot{k}_0(t) = k_{11}(t)h_1 + k_{22}(t)h_2$$

$$\cdot \; -\frac{1}{2}\begin{bmatrix} s_1 & s_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\frac{1}{r+r_1(t)} - \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}\frac{1}{s-s_1(t)} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$$

$$= k_{11}(t)h_1 + k_{22}(t)h_2 -$$

$$\frac{1}{2}\begin{bmatrix} s_1 & s_2 \end{bmatrix} \begin{bmatrix} r-s+r_1(t)+s_1(t) & r-2s+r_1(t)+2s_1(t) \\ r-2s+r_1(t)+2s_1(t) & r-4s+r_1(t)+4s_1(t) \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$$

$$-\frac{d}{dt}\begin{bmatrix} s_1(t) \\ s_2(t) \end{bmatrix} = \begin{bmatrix} 1+r+r_1(t)-s+s_1(t) & r-2s+r_1(t)+2s_1(t) \\ 1+r-2s+r_1(t)+2s_1(t) & 2+r-4s+r_1(t)+4s_1(t) \end{bmatrix}\begin{bmatrix} s_1(t) \\ s_2(t) \end{bmatrix}$$

$$-\begin{bmatrix} \dot{k}_{11}(t) & \dot{k}_{12}(t) \\ \dot{k}_{12}(t) & \dot{k}_{22}(t) \end{bmatrix}$$

$$= \begin{bmatrix} k_{11}(t) & k_{12}(t) \\ k_{12}(t) & k_{22}(t) \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} k_{11}(t) & k_{12}(t) \\ k_{12}(t) & k_{12}(t) \end{bmatrix} +$$

$$\begin{bmatrix} k_{11}(t) & k_{12}(t) \\ k_{12}(t) & k_{22}(t) \end{bmatrix}\begin{bmatrix} r-s+r_1(t)+s_1(t) & r-2s+r_1(t)+2s_1(t) \\ r-2s+r_1(t)+2s_1(t) & r-4s+r_1(t)+4s_1(t) \end{bmatrix} x$$

$$\begin{bmatrix} k_{11}(t) & k_{12}(t) \\ k_{12}(t) & k_{22}(t) \end{bmatrix} \tag{5.104}$$

$$\text{with} \quad \begin{matrix} k_0(T) = 0 & s_1(T) = 0 & s_2(T) = 0 \\ k_{12}(T) = 0 & k_{11}(T) = f_1 & k_{22}(T) = f_2 \end{matrix} \tag{5.105}$$

Having illustrated the solution of the stochastic game with complete information we next consider the random partial information case. The derivations in these sections follow |63-64|.

(b) Random Partial Information : In order to illustrate this case we consider the following game. Given that

$$dx(t) = A(t)x(t)dt + B(t)u(t)dt + C(t)v(t)dt + d\eta(t) \qquad (5.106)$$
$$dy(t) = H_1 dx(t) + d\xi(t) \qquad (5.107)$$
$$dz(t) = H_2 dx(t) + d\zeta(t) \qquad (5.108)$$

represent the system and observation equations of the players where the quantities $x(t)$, $u(t)$, $v(t)$, $d\eta(t)$, $A(t)$, $B(t)$, $C(t)$ are used as in the preceding section (a) and $y(t)$ is an m-vector of observations of player I, $z(t)$ is an l-vector of observations of player II, and the constraints on u, v are given by

$$U = \left\{ u_i(t) : |u_i(t)| \leq L < \infty \text{, for each } t \in [t_0, T], i = 1, \ldots, r \right\} \qquad (5.109)$$

$$V = \left\{ v_i(t) : |v_i(t)| \leq L < \infty \text{, for each } t \in [t_0, T], i = 1, \ldots, s \right\} \qquad (5.110)$$

where L is a large enough number. The payoff function to player II from player I is given by

$$I(u,v) = E_\eta \left\{ \int_0^T (\langle u(t), Ru(t)\rangle - \langle v(t), Sv(t)\rangle)dt + \langle x(T), Fx(T)\rangle \right\} \qquad (5.111)$$

We are required to find optimal control strategies for the players satisfying assumption 2.1 in chapter 2, and is such that it is immaterial which player chooses his strategy first. The saddlepoint inequality (5.62) as viewed by the two players can be split into

$$\min_{u \in U} I(u, v^0) = I(u^0, v^0) \qquad (5.112)$$

for player I and

$$\max_{v \in V} I(u^0, v) = I(u^0, v^0) \qquad (5.113)$$

for player II.

We first consider the problem in Eq. (5.112) for player I, who must base his strategy on the observation set

$$Y_t = \{y(s): 0 \leq s \leq t, \; dy(s) = H_1 dx(s) + d\xi(s)\} \qquad (5.114)$$

Then instead of (5.112) player I can verify

$$E_{Y_t} \; I(u^0, v^0) \leq E_{Y_t} \; I(u, v^0) \qquad (5.115)$$

where E_{Y_t} stands for the conditional expectation given Y_t at time $t \in [t_0, T]$.

Let us consider the problem as starting at (x_0, t_0). By the imbedding and optimality principles of dynamic programming we have

$$E_{Y_{t_0}} \; E_\eta \left\{ \int_{t_0}^{T} (\langle u^0(t), Ru^0(t)\rangle - \langle v^0(t), Sv^0(t)\rangle) dt + \langle x^0(T), Fx^0(T)\rangle \right\}$$

$$\leq E_{Y_{t_0}} \; E_\eta \left\{ \int_{t_0}^{T} (\langle u(t), Ru(t)\rangle - \langle v^0(t), Sv^0(t)\rangle) dt + \langle x^0(T), Fx^0(T)\rangle \right\}$$

$$(5.116)$$

Let $E_{Y_{t_0}} \underline{W}(x_0, t_0)$ be defined to be the left hand side of (5.116) then

$$E_{Y_{t_0}} \; \underline{W}(x_0, t_0)$$

$$\leq E_{Y_{t_0}} \; E_\eta \left\{ \int_{t_0}^{t_0+dt_0} (\langle u(t_0), Ru(t_0)\rangle - \langle v^0(t_0), Sv^0(t_0)\rangle) dt_0 + \underline{W}(x_0 + dx_0, t_0 + dt_0) \right\}$$

$$= E_{Y_{t_0}} E \quad ((\langle u(t_0), Ru(t_0)\rangle - \langle v^0(t_0), Sv^0(t_0)\rangle)\, dt_0 + \underline{W}(x_0,t_0)$$

$$+\langle \frac{\partial \underline{W}}{\partial x}, dx\rangle + \frac{\partial \underline{W}}{\partial t} dt + \frac{1}{2} \langle dx, \frac{\partial^2 \underline{W}}{\partial x^2} dx\rangle \qquad (5.117)$$

Hence

$$- E_{Y_{t_0}} \frac{\partial \underline{W}\ (x_0,t_0)}{\partial t}$$

$$\leq E_{Y_{t_0}} \quad \langle u(t\ ), Ru(t\ )\rangle - \langle v(t\ ), Sv(t\ )\rangle + \frac{1}{2} \mathrm{tr}(\frac{\partial^2 \underline{W}}{\partial x^2} S)$$

$$+ \langle \frac{\partial \underline{W}}{\partial x}, A(t)x(t) + B(t)u(t) + C(t)v(t)\rangle \qquad (5.118)$$

Equality is obtained for $u = u^0$ where

$$u^0(t_0) = -R^{-1} B^T(t_0)\, E_{Y_{t_0}} \frac{\partial \underline{W}}{\partial x_0} \qquad (5.119)$$

Let $t_0 \to t$, $x_0 \to x$. Henceforth we can drop the subscript o. Substituting Eq. (5.119) in (5.118) we have

$$- E_{Y_t} \frac{\partial \underline{W}(x,t)}{\partial t}$$

$$= -\frac{1}{2} \langle E_{Y_t} \frac{\partial \underline{W}(x,t)}{\partial x}, B(t)R^{-1} B^T(t)\, E_{Y_t} \frac{\partial \underline{W}(x,t)}{\partial x} \rangle - \langle v^0(t), Sv^0(t)\rangle$$

$$+ E_{Y_t} \langle \frac{\partial \underline{W}(x,t)}{\partial x}, A(t)x(t)+C(t)v^0(t)\rangle + \frac{1}{2} \mathrm{tr}\left\{E_{Y_t} \frac{\partial^2 \underline{W}(x,t)}{\partial x^2}\ S\right\}$$

$$(5.120)$$

Let us consider a series solution

$$\underline{W}(x,t) = \underline{k}_0(t) + \langle \underline{g}(t), x\rangle + \langle x, \underline{K}(t)x\rangle \qquad (5.121)$$

which is separable in x and t.

Let

$$\underline{m}(t) = E_{Y_t}(x) \qquad \underline{\Sigma}(t) = \mathrm{Var}_{Y_t}(x) \qquad (5.122)$$

Now consider

$$E_{Y_t}\frac{\partial W(\underline{x},t)}{\partial t} = \dot{\underline{k}}_0(t)+\langle\dot{\underline{g}}(t),\underline{m}(t)\rangle+\langle\underline{m}(t),\dot{\underline{K}}(t)\underline{m}(t)\rangle+\ \mathrm{tr}\left\{\dot{\underline{K}}(t)\mathrm{Var}_{Y_t}(x)\right\} \tag{5.123}$$

$$E_{Y_t}\frac{\partial W(\underline{x},t)}{\partial x} = \underline{g}(t) + 2\underline{K}(t)\ \underline{m}(t) \tag{5.124}$$

$$E_{Y_t}\frac{\partial^2 W(\underline{x},t)}{\partial x^2} = \underline{K}(t) \tag{5.125}$$

Substituting (5.123 - 5.125) into (5.120) we obtain on equating the coefficients,

$$-\dot{\underline{k}}_0(t) = -\frac{1}{2}\langle\underline{g}(t),B(t)R^{-1}B^T(t)\underline{g}(t)\rangle + \frac{1}{2}\,\mathrm{tr}\left\{\underline{K}(t)G\right\}$$

$$+\ \langle\underline{g}(t),C(t)v^0(t)\rangle - \langle v^0(t),Sv^0(t)\rangle \tag{5.126}$$

$$-\dot{\underline{g}}(t) = -2\underline{K}(t)\ B(t)R^{-1}\ B^T(t)\underline{g}(t) + A^T(t)\underline{g}(t)$$

$$+\ 2\underline{K}(t)\ C(t)v^0(t) \tag{5.127}$$

$$-\dot{\underline{K}}(t) = \underline{K}(t)\ A(t) + A^T(t)\underline{K}(t) - 2\underline{K}(t)B(t)R^{-1}B^T(t)\underline{K}(t) \tag{5.128}$$

with the boundary conditions

$$\underline{g}(T) = 0\ ,\quad \underline{k}_0(T) = 0\ ,\quad \underline{K}(T) = F \tag{5.129}$$

The optimal control strategy for player I is then

$$u^0 = -R^{-1}B^T(t)\ (\ \underline{g}(t) + 2\underline{K}(t)\underline{m}\ ) \tag{5.130}$$

To determine $\underline{m}(t)$ some other criterion has to be invoked by player I. Let us say that he utilises the supercriterion that $\underline{m}(t)$ be such as to minimise

$$\mathrm{tr}\ E_{Y_t}\left\{\ (\ x(t)-\underline{m}(t)(\ (x(t)-\underline{m}(t)\ )^T\right\} = \mathrm{tr}\left\{\underline{P}(t)\right\} \tag{5.131}$$

With this the player makes use of the separation theorem under the assumption that player II plays optimally. The solution to this problem is well known [65] and the corresponding solutions are summarised below.

Filter Equations for Player I:

$$\dot{\underline{m}}(t) = A(t)\underline{m}(t) + B(t)u^{o}(t) + C(t)v^{o}(t) + \underline{G}(t)(y(t) - H_1\underline{m}(t)) \qquad (5.132)$$

$$\underline{G}(t) = \underline{P}(t)\, H_1^T \Xi^{-1} \qquad (5.133)$$

Variance Equations for Player II:

$$\dot{\underline{P}}(t) = \theta - \underline{P}(t)H_1^T \Xi^{-1} H_1\underline{P}(t) + (A(t) - 2B(t)R^{-1}B^T(t)K(t))\underline{P}(t) + \underline{P}(t)\,(A(t) - 2B(t)R^{-1}B^T(t)K(t))^T \qquad (5.134)$$

$$\underline{P}(0) = \mathrm{cov}\,(x(t_o)\ \ x(t_o)^T) \qquad (5.135)$$

In a similar manner, the optimal control strategy, the filter and variance equations can be derived for player II by considering instead Eq. (5.113) and requiring the maximisation conditioned on

$$Z_t = z(s):\ 0 \le s \le t\,, \quad dz(s) = H_2 dx(s) + d\,(s) \qquad (5.136)$$

The results are obtained from changing under-bars to over-bars in the equations for player I and

$$B(t) \rightarrow C(t), \quad H_1 \rightarrow H_2\,, \quad \underline{P} \rightarrow \bar{P}\,, \quad R^{-1} \rightarrow -S^{-1}$$
$$\underline{G} \rightarrow \bar{G}\,, \quad \Xi \rightarrow \Im\,, \quad \underline{m} \rightarrow \bar{m} \qquad (5.137)$$

Hence we have

Value Equation:

$$E_{Z_t} \bar{W}(x,t) = \bar{k}_0(t) + \langle \bar{T}(t), \hat{x} \rangle + \langle \hat{x}, \bar{K}(t)\, \hat{x} \rangle + tr \left\{ \bar{K}(t)\, \bar{\Sigma} \right\} \qquad (5.138)$$

$$\bar{\Sigma} = Var\; E_{Z_t}(x) \qquad (5.139)$$

Optimal Control Strategy:

$$v^0 = S^{-1} C^T(t) \left(\bar{T}(t) + 2\bar{K}(t)\hat{x} \right) \qquad (5.140)$$

Filter Equations for Player II:

$$\dot{\hat{x}}(t) = A(t)\hat{x}(t) + B(t)u^0 + C(t)v^0 + \bar{G}(t) \left(z(t) - H_2\hat{x}(t) \right) \qquad (5.141)$$

where

$$\bar{G}(t) = \bar{P}(t)\, H_2^T\, 3^{-1} \qquad (5.142)$$

Variance Equations for Player II:

$$\dot{\bar{P}}(t) = S(t) - \bar{P}(t) H_2^T 3^{-1} H_2 \bar{P}(t) + \left(A(t) + 2C(t)S^{-1}C^T(t)\bar{K}(t) \right) \bar{P}(t)$$

$$+ \bar{P}(t) \left(A(t) + 2C(t)S^{-1}C^T(t)\bar{K}(t) \right)^T \qquad (5.143)$$

$$\bar{P}(t) = cov \left((x(t) - \hat{x}(t))\, (x(t) - \hat{x}(t))^T \right) \qquad (5.144)$$

$$\bar{P}(0) = cov \left(x(t_0)\;\; x(t_0)^T \right) \qquad (5.145)$$

Thus the moment we had to consider a game with random partial information, the players had to invoke a supercriterion to determine the position of the game and make the assumption that the other player uses an optimal control strategy. Since the supercriterion used is subjective the game is to be looked

upon as equivalent to a game with two criteria for a player which are hierarchical. We used both supercriteria for the players as the minimum variance estimators. Nothing prevents one of the players to use a maximum likelihood estimator. The game now settles to a Nash equilibrium point since <u>each player assumes the other player to play optimally.</u> We can compare the results obtained here with the partial information case in chapter 3. With random partial information, we could not reduce the game completely to an equivalent two-person zero-sum game with complete information and thus assure a saddlepoint for the specified single payoff function.

5.5 CONCLUSIONS

We conclude here our initial investigations of games with uncertainty. To our delight, we find a vast as yet unexplored area wherein the theory of stochastic differential games can be suitably applied. A number of conceptual problems have been raised and some of these were demonstrated for the game with random partial information. The concept of incomplete structural information will be examined in more detail in chapter 7. In the next chapter we consider the Markov Positional Game which is related to the Extensive Game of Kuhn.

VI MARKOV POSITIONAL GAMES

We describe herein the discrete-time Markov Positional Game model, which, as we have already remarked in chapter 1, is analogous to the dual control problem. Markov Games have been studied by Zachrisson |15|, who considers many game theoretic aspects of a similar decision-making model due to Howard |16|. Markov Games can be imbedded in the extensive game structure of Kuhn |9| or of Aumann |66|. We shall follow this procedure to imbed the Markov Positional Game in the extensive game to study its properties. Other generalisations of the one-sided decision-making model are due to Karp and Hoffman |67|.

In the last chapter we have raised many conceptual problems. (See also |15|). Making use of Bayesian methods (as in the theory of dual control), we show the existence of a different state space for each player and the umpire. We then consider the relevance of different types of strategies. Finally we give algorithms to determine the optimal control strategies.

6.2 DESCRIPTION OF THE GAME

We are given that the set of equations

$$x_{k+1} = g_k(x_k, u_k, v_k, \eta_k) \tag{6.1}$$

$$y_k = h_k^1 (x_k, \xi_k) \tag{6.2}$$

$$z_k = h_k^2 (x_k, \zeta_k) \tag{6.3}$$

constitute the game with constraints involving inputs and observations, positions and observations being assumed to be incorporated in (6.1 - 6.3) and the sets U_k, V_k. Here

$x_k \in X_k \subseteq \mathcal{R}^n$ is the position governing the physical evolution of the game.

$u_k \in U_k \subseteq \mathcal{R}^r$ is the control action of player I.

$v_k \in V_k \subseteq \mathcal{R}^s$ is the control action of player II.

$y_k \in Y_k \subseteq \mathcal{R}^m$ is the observation vector of player I.

$z_k \in Z_k \subseteq \mathcal{R}^l$ is the observation vector of player II.

$\eta_k, \xi_k, \zeta_k, \in W_q$ are chance variables of the game (noise processes).

Let W_1 denote the set of all randomisations of elements in $U = \bigcup_k U_k$. Hence $\tilde{u} \in U \times W_1 \triangleq \tilde{U}$ constitutes a (possible) mixing of strategies for player I. (The tilde denotes a mixed strategy). Let W_2 denote the set of all randomisations of elements in $V = \bigcup_k V_k$. Hence $\tilde{v} \in V \times W_2 \triangleq \tilde{V}$, constitutes a (possible) mixing of strategies.

The termination of the game is specified thus: The goal of the players is to transfer the position of the game given by (6.1) through observations constituted by (6.2 - 6.3) starting at x_0 (the initial position) specified to both the players, to the terminal position x_N specified to both the players. The payoff to player II from player I is given by

$$I(u,v) = E\left(\Sigma_{k=1}^{N} L_k(x_k, u_{k-1}, v_{k-1}) \right) \tag{6.4}$$

where E denotes the expectation operator over all randomisations.

Definition 6.1 : Let $\tilde{u} \in \tilde{U}_0 \subseteq \tilde{U}$, $\tilde{v} \in \tilde{V}_0 \subseteq \tilde{V}$ constitute a pair of strategies $(\tilde{u},\tilde{v})$ such that the players can assure the termination of the game in the prescribed sense. Let this be true for arbitrary $\tilde{u} \in \tilde{U}_0$ and $\tilde{v} \in \tilde{V}_0$. Then we say that $(\tilde{u},\tilde{v})$ constitutes a _playable pair of mixed strategies_.

The above game can now be written in terms of playable pair strategies as $(X , U_0, V_0, Y , Z, W, S, H^1, H^2, I, \Theta)$ where $X = \bigcup_k X_k$, $Y = \bigcup_k Y_k$, $Z = \bigcup_k Z_k$, $W = W_1 \times W_2 \times W_q$, $H^1 = \bigcup_k H^1_k$, $H^2 = \bigcup_k H^2_k$ and Θ is the interval over which the game is played. We shall henceforth denote this game by (U_0,V_0,I). In this game we are required to find a saddlepoint for the payoff function $I(\tilde{u},\tilde{v})$ in mixed strategies.

$$I(\tilde{u}^*,\tilde{v}) \leq I(\tilde{u}^*,\tilde{v}^*) \leq I(\tilde{u},\tilde{v}^*) \tag{6.5}$$

To determine (6.5) we need an understanding of mixed strategies in a sequential game such as (U_0,V_0,I). We further require methods to determine optimal strategies and hence the value of the game. This is done by imbedding the game (U_0,V_0,I) into an equivalent Markov Positional Game.

6.3 STATE DESCRIPTION OF THE POSITIONAL GAME

It is necessary to convert the positional game into a Markov Positional Game so that the Markovian properties of the conditional probability densities can be used to define the various strategies. We give a method to construct the state space of this equivalent Markov Game. Since this is similar to the method of constructing the augmented state

vector of a stochastic system as given by Aoki |60|, we shall discuss the method only in its game context.

Suppose player II has chosen a priori his optimal mixed strategy $v^* \in V_0$. Then the problem as seen by player I is essentially that given in |60|. To him, the stochastic processes v_k^*, η_k, ξ_k all constitute noise processes. The player is to base his strategy on the observations $y^k = \{y_1, y_2, \ldots, y_k\}$ and his past control actions $u^{k-1} = \{u_1, u_2, \ldots, u_{k-1}\}$. Since these are growing vectors, and since he has to determine the position in the presence of unknown noise processes, he may choose to infer these processes from (u^{k-1}, y^k) through data processing. Further both the players have the same a priori specifications of the noise processes.

Let the player I construct the following state vector. (See chapter 2).

$$\phi_k = (x_k, v^{k-1}, \eta^k, \xi^k, y^k, u^{k-1}) \qquad (6.6)$$

This has a lot of redundant information. We consider a reduced state vector for player I. Let s_k be a vector whose dimension is still to be specified, such that

$$\phi_k' = (x_k, s_k, y_k, u_{k-1}) \qquad (6.7)$$

is an equivalent state vector in the sense that

$$\text{Prob.}(\phi_k) = \text{Prob.}(\phi_k') \qquad (6.8)$$

and

$$\begin{aligned} \text{Prob.}(\phi_k|\phi^{k-1}) &= \text{Prob.}(\phi_k|\phi_{k-1}) \\ = \text{Prob.}(\phi_k'|\phi'^{k-1}) &= \text{Prob.}(\phi_k'|\phi_{k-1}') \end{aligned} \qquad (6.9)$$

Thus ϕ'_k has the Markovian property we desire. Player I can now choose his strategy based on (s_k, y_k, u_{k-1}), since he does not know x_k. The vector s_k satisfies the recurrence equation

$$s_{k+1} = \Sigma_k(s_k, y_{k+1}, u_k) \tag{6.10}$$

and the optimal strategy for player I can be written

$$u_k = U_k(s_k, y_k, w_1) \tag{6.11}$$

where $w_1 \in W_1$, thus including a possible mixed strategy. The existence of such a vector s_k of finite dimension is intimately related with the property of the distribution function governing ϕ_k, as has been shown by Aoki [68]. Thus, in order that a player have a reduced state vector of finite dimensions, the distribution functions must be of the self-reproducing type which assures a sufficient statistic for the estimation process. The sufficient statistics satisfies (6.10) and hence is a candidate for the reduced state vector in (6.7). Where this is not possible, the conditional density is itself a candidate. The state space is not of finite dimension now.

Next, consider the game as viewed by player II, when he is told that player I has chosen an optimal mixed strategy $\tilde{u}^*$. A similar process of reasoning leads to a reduced state space for player II.

$$\psi_k - \psi'_k = (z_k, s'_k, s_k, v_{k-1}) \tag{6.12}$$

where s'_k is the function yet to be determined in a like manner.

Consider then the vector

$$\chi_k = \phi_k \cup \psi_k = (x_k, s_k, s'_k, z_k, y_k, u_{k-1}, v_{k-1}) \tag{6.13}$$

This is the vector describing the state of the game as seen by an independent observer whom we appropriately term the 'umpire'. We can partition the vector χ_k into subclasses consisting of vectors <u>observed</u>, <u>inferred</u>, <u>remembered</u> and <u>ignored</u> by a player.

χ_k = observed state $\oplus$ inferred state $\oplus$ remembered state $\oplus$ ignored state.

Each such partitioning differs from player to player. We summarize the above in the theorem.

<u>Theorem</u> 6.1 : <u>Given a discrete-time two-person zero-sum positional game as described above, a state space for the game can be constructed which is the union of the state spaces of the players in the game. Each player views the state space in four disjoint classes, of states observed, inferred, remembered and ignored. The partitioning of the state space may differ from player to player. Further the state space has finite dimensions if and only if the conditional probability distributions of every player has sufficient statistics.</u>

We thus see that Zachrisson's conjecture |15| on partitioning of the state space holds in our game. Since we have imbedded the given positional game into another positional game whose state vector satisfies the Markov property, we term the latter game the <u>Markov Positional Game</u>. Therefore, we

restrict ourselves to the following problem. Given

$$z_{k+1} = s_k \ (z_k, u_k, v_k) \qquad (6.14)$$

$$y_k = \nu_k \qquad (6.15)$$

$$w_k = \rho_k \qquad (6.16)$$

where

$$x_k = \mu_k \ \nu_k \ \rho_k \qquad (6.17)$$

$$\mu_k \cap \nu_k = \emptyset, \mu_k \cap \rho_k = \emptyset \qquad (6.18)$$

with Prob. (z_0) specified by the problem, find the sequences ($\{\tilde{u}_k^*\}$, $\{\tilde{v}_k^*\}$) of optimal mixed strategies such that the payoff function from player I to player II is given by

$$I(\tilde{u}^*, \tilde{v}^*) = E\left(\Sigma_{k=1}^{N} L_k(x_k, u_k, v_k) \right) \quad x_k \leq \mu_k \qquad (6.19)$$

has a saddlepoint as in (6.8).

6.4 PROPERTIES OF DIFFERENT TYPES OF STRATEGIES

We now enquire into the properties of various subsets of the set of all mixed strategies for a player in the Markov Positional Game. Intuitively, a pure strategy of a player involves no randomisation, i.e., he has only randomised strategies with the Dirac delta function as the density function on his information set. The notions of an extensive game (due to Aumann) tells us how the action and information spaces should be chosen. We give the relevant definitions in Appendix C.

Since the state vector ϕ_k or ψ_k forms part of the state vector χ_k of the game, he can, in principle, determine the state of the other player due to the Markovian property. Thus to each player $\phi = \phi_1 \times \dots \times \phi_k$, $(\psi = \psi_1 \times \dots \times \psi_k)$ constitutes a sequence of information spaces, while the corresponding sequence of his action space is $U_o = U_{1o} \times \dots \times U_{ko}$ $(V_o = V_{1o} \times \dots \times V_{ko})$. Thus from an individual player's point of view, the game been normalised for the other player. If we now consider a game of perfect recall, a player can unwind the sequences of information patterns he had for these choices right up to the start of the game. He will also be able to know the similar sequence for the other player. The sequence of transformations required to lead to a game of perfect recall (see Appendix D) are assured by the very Markovian nature of the state. Then we are guaranteed by Kuhn's theorem of the existence of a behaviour strategy for a player when he has perfect recall. In reference |66|, the interrelation between mixed and behaviour strategies has been discussed in detail. The nature of randomisation in a general mixed strategy being quite involved, we seek to simplify the process of randomisation by imposing restrictions. We give below the various definitions of different strategies.

<u>Definition 6.2</u> : A <u>mixed strategy</u> is a sequence of (measurable) transformations $u = (u_1, u_2, \dots)$, $u_i : \Omega \times \phi \rightarrow U_i$ where $\Omega \subset W_1$ is a fixed sample space.

Definition 6.3 : A behaviour strategy is a mixed strategy b such that for $i \neq j$, $b_i(\cdot,\phi_i)$ and $b_j(\cdot,\phi_j)$ are mutually independent random variables. $\phi_i \in \Phi_i$ and $\phi_j \in \Phi_j$.

Definition 6.4 : A pure strategy is a degenerate mixed strategy which assigns probability 1 to a single control function u (say) and zero to others.

Definition 6.5 : A stationary strategy is a behaviour strategy $b_i(\cdot,\phi)$ such that for a given i $b_i(\cdot,\phi)$ is a stationary stochastic process along the ϕ axes. |67|.

Definition 6.6 : Two mixed strategies are said to be equivalent if they determine the same distributions on U.

The above definitions were independent of the payoff functions, while the next two are defined in terms of the payoff functions.

Definition 6.7 : Let $I(\tilde{u},\tilde{v})$ denote the payoff due to the playable pair $(\tilde{u},\tilde{v})$. Then a strategy $\tilde{u}^* \in \tilde{U}_o$ is said to be Bayes with respect to v if

$$I(\tilde{u}^*,\tilde{v}) = \inf_{\tilde{u}^* \in \tilde{U}_o} I(\tilde{u},\tilde{v}) \qquad (6.20)$$

Definition 6.8 : If $(\tilde{u}^*,\tilde{v}^*)$ is a playable pair such that $(\tilde{u}^*,\tilde{v}^*)$ satisfies

$$I(\tilde{u}^*,\tilde{v}^*) = \inf_{\tilde{u} \in \tilde{U}_o} \sup_{\tilde{v} \in \tilde{V}_o} I(\tilde{u},\tilde{v}) \qquad (6.21)$$

then the strategy $\tilde{u}^*,\tilde{v}^*$ is the inf-sup strategy for the players.

Clearly there exists a relation between the Bayes and inf-sup strategies. This is discussed by Sworder|23|.

Definition 6.9 : If for every $\epsilon > 0$ there exists a $\tilde{v}_\epsilon \in \tilde{V}_o$ such that

$$I(\tilde{u}^*,\tilde{v}_\epsilon) \leq \inf_{\tilde{u}^* \in \tilde{U}_o} I(\tilde{u},\tilde{v}_\epsilon) + \epsilon \qquad (6.22)$$

then $\tilde{u}^*$ is called an extended Bayes strategy.

Definition 6.10 : If there exists a $\tilde{u}^* \in \tilde{U}_o$ such that

$$I(\tilde{u}^*,\tilde{v}) = C \qquad (6.23)$$

where C is a fixed constant, for all $\tilde{v} \in \tilde{V}_o$, $\tilde{u}^*$ is called an equaliser strategy.

We have now two important theorems linking some of these strategies.

Theorem 6.2 : (Derman)|69| : From the viewpoint of one player the normalised one-player Markovian Positional Game has a stationary strategy in $\tilde{u}^o \in \tilde{U}_B$, the set of all behaviour strategies in $\tilde{U}_o$ such that

$$I(\tilde{u}^o,\tilde{v}^{**}) = \min_{\tilde{u}^o \in \tilde{U}_B} I(\tilde{u},\tilde{v}) \qquad (6.24)$$

Theorem 6.3 : If $\tilde{u}^* \in \tilde{U}_o$ is an equaliser strategy and is extended Bayes, then it is also a inf-sup strategy and the game has a value|23|.

Thus we are now led to consider stationary strategies which are also equalizer strategies. Combining theorems 6.2 and 6.3 we have the following theorem.

<u>Theorem 6.4</u> : Let $(\tilde{u}^*, \tilde{v}^*)$ constitute a playable pair of equalizer strategies of the players. If $(\tilde{u}^*, \tilde{v}^*)$ are also stationary strategies and further are extended Bayes for each player then the game (U_0, V_0, I) has a saddlepoint in mixed strategies where only equaliser strategies are used.

We thus see that a Markovian Positional Game can have equalizer strategies that are also stationary. In fact if a strategy demands a growing vector ν^k on which to base itself, the Markovian nature is destroyed.

6.5 OPTIMAL STRATEGIES FOR PLAYER I

Having seen the connection of the Markov Positional Game with the extensive games, we now determine explicit algorithms for the players to choose their optimal strategies for the positional game. The umpire's state x_k constitutes a first-order Markov process with known transition probability densities. Similar to the game in (6.14 - 6.18), we consider strategies determined based only on ϕ_k , the known state vector of player I. We determine the strategy for player I assuming player II to have chosen his optimal strategy first.

(a) Last Stage : Let us determine $\tilde{u}^*_{N-1}$, assuming $\tilde{u}^*_0, \ldots, \tilde{u}^*_{N-2}$. Let us assume that player II has already chosen his optimal strategy. Then the only term to minimise is

$$E\,(L_M(x_M,u_{M-1},v^*_{M-1})|\nu^{M-1},\rho^{M-1})$$

$$= \int L_M(x_M,u_{M-1},v_{M-1})\,p(x_M,u_{M-1},v_{M-1}|\nu^{M-1},\rho^{M-1})\,d(x_M,u_{M-1},v_{M-1})$$

$$= \int L_M(x_M,u_{M-1},v_{M-1})\,p(x_M|u_{M-1},v_{M-1},\ {}_{M-1},\ {}_{M-1})$$

$$\times\ p(u_{M-1},v_{M-1}|\ {}^{M-1},\ {}^{M-1})\,d(x_M,u_{M-1},v_{M-1})$$

$$= \int L_M(x_M,u_{M-1},v_{M-1})\,p(\mu_M|\mu_{M-1},u_{M-1},v_{M-1},\nu_{M-1},\rho_{M-1})$$

$$\times\ p(\mu_{M-1}|u_{M-1},v_{M-1},\nu^{M-1},\rho^{M-1})\,p(u_{M-1}|v_{M-1},\nu^{M-1})$$

$$\times\ p(v_{M-1}|\rho^{M-1})\,d(\mu_M,\mu_{M-1},u_{M-1},v_{M-1}) \qquad (6.25)$$

Since u_{M-1} is independent of ρ^{M-1} and v_{M-1} is independent of ν^{M-1} and $p(v_{M-1}|\rho^{M-1})$ is chosen optimally by player II, we have the left hand side of (6.25)

$$= \int \lambda_M\,p(u_{M-1}|v^*_{M-1},\nu^{M-1})\,p^*(v_{M-1}|\rho^{M-1})\,d(u_{M-1},v_{M-1}) \qquad (6.26)$$

where

$$\lambda_M = \int L_M(x_M,u_{M-1},v_{M-1})\,p(\mu_M|\mu_{M-1},u_{M-1},v_{M-1},\nu_{M-1},\rho_{M-1})$$

$$\times\ p(\mu_{M-1}|u_{M-1},v_{M-1},\nu^{M-1},\rho^{M-1})\,d(\mu_M,\mu_{M-1}) \qquad (6.27)$$

and

$$d(x,y) = dxdy \qquad (6.28)$$

assuming $p(\mu_{M-1}|u_{M-1},v_{M-1},\nu^{M-1},\rho^{M-1})$ is available and $p(\mu_M|\mu_{M-1},u_{M-1},v_{M-1},\nu^{M-1},\rho^{M-1})$ is computable from the known

transition densities. Then a mixed strategy is the determination of the distribution $\tilde{u}^*_{M-1} = p(u_{M-1}|\tilde{v}^*_{M-1}, \nu^{M-1})$ for player I, player II having chosen his mixed strategy $\tilde{v}^*_{M-1} = p(v_{M-1}|\rho^{M-1})$. Since we have considered a game of perfect recall, we can use behaviour strategies. Hence player I need determine only $p(u_{M-1}|\tilde{v}^*_{M-1}, \nu_{M-1})$ and player II need determine $p(v_{M-1}|\rho_{M-1})$ which are randomised, based on the information available at the (M-1)-th instant only. Denote

$b_{u_{M-1}} = b_{u_{M-1}|\nu_{M-1}, \tilde{v}^*_{M-1}}$ = behaviour strategy for player I at (M-1)-th stage.

$b_{v_{M-1}} = b_{v_{M-1}|\rho_{M-1}}$ = behaviour strategy for player II at (M-1)-th stage

Let $b^s_{u_{M-1}}$, $b^s_{v_{M-1}}$ denote the corresponding stationary strategies; then we can write

$$\gamma^*_M = \min_{b^s_{u_{M-1}}} \max_{b^s_{v_{M-1}}} \lambda_M \tag{6.29}$$

where λ_M is the nonoptimal value for both players. Again, by theorem 6.4, if these also happen to be equaliser strategies and are also extended Bayes for each player then we also have

$$\gamma^*_M = \max_{b^s_{v_{M-1}}} \min_{b^s_{u_{M-1}}} \lambda_M \tag{6.30}$$

where λ_M is the nonoptimal value for both players.

(b) <u>Last 2 Stages</u> : Let us now determine the optimal mixed strategies at the last but one stage, such that $p(u_{M-2}|\tilde{v}^*_{M-2}, \nu^{M-2})$

followed by $p^*(u_{N-1}|\tilde{v}^*_{N-1},\nu^{N-1})$ minimises $E(L_{N-1}+L_N|\nu^{N-2})$, player II having chosen his mixed strategy optimally. We again have

$$E(L_{N-1}|\nu^{N-2}) = \int\lambda_{N-1}p(u_{N-2}|v^*_{N-2},\nu^{N-2})p(v_{N-2}|\rho^{N-2})d(u_{N-2},v_{N-2}) \tag{6.31}$$

where λ_{N-1} is defined in the obvious way and we assume $p(\mu_{N-1}|\mu_{N-2},u_{N-2},v_{N-2},\nu_{N-2},\rho_{N-2})$ is computable from the known transition probability density $d(\chi_{N-1}|\chi_{N-2})$. Now

$$E(L_N|\nu^{N-2}) = E(\,E(L_N|\nu^{N-1})|\nu^{N-2}) \tag{6.32}$$

Hence

$$\begin{aligned}\min_{\tilde{u}^*_{N-2}} \quad & E(\,L_N + L_{N-1}|\nu^{N-2}\,)_{\tilde{u}^*_{N-1}} \\ = \quad & \int\lambda_{N-1}p(u_{N-2}|\tilde{v}^*_{N-2},{}^{N-2})p(v_{N-2}|\nu^{N-2})d(u_{N-2},v_{N-2}) \\ & + \int\gamma^*_N\, p(\nu^{N-1},\rho^{N-1}|u_{N-2},v_{N-2},\nu^{N-2},\rho^{N-2})p(u_{N-2}|\nu^{N-2},v_{N-2}) \\ & \qquad \times \quad p(v_{N-2}|\rho^{N-2})d(\nu_{N-1},\rho_{N-1},u_{N-2},v_{N-2})\end{aligned} \tag{6.33}$$

Define

$$\gamma_{N-1} = \lambda_{N-1} + \int\gamma^*_N\, p(\nu_{N-1},\rho_{N-1}|u_{N-2},v_{N-2},\nu^{N-2},\rho^{N-2})d(\nu_{N-1},\rho_{N-1}) \tag{6.34}$$

where again we assume $p(\nu_{N-1},\rho_{N-1}|u_{N-2},v_{N-2},\nu^{N-2},\rho^{N-2})$ is is known. $\tilde{u}^*_{N-2}$ is found by

$$\gamma^*_{N-1} = \min_{\tilde{u}^*_{N-2}} \gamma_{N-1} \tag{6.35}$$

Again, using stationary strategies and invoking theorem 6.4 we have

$$\gamma^*_{N-1} = \min_{b^s_{u_{N-2}}} \max_{b^s_{v_{N-2}}} \gamma_{N-1} = \max_{b^s_{v_{N-2}}} \min_{b^s_{u_{N-2}}} \gamma_{N-1} \qquad (6.36)$$

(c) <u>In General for k Stages</u> : Define

$$\lambda_k = \int \Omega_k(x_k, u_{k-1}, v_{k-1}) p(\mu_k | \mu_{k-1}, \nu_{k-1}, \rho_{k-1}, u_{k-1}, v_{k-1})$$

$$\times \; p(\mu_{k-1} | u_{k-1}, v_{k-1}, \nu^{k-1}, \rho^{k-1}) d(\mu_k, \mu_{k-1}) \qquad (6.37)$$

and

$$\gamma_k = \lambda_k + \int \gamma^*_{k+1} \, p(\nu_k, \rho_k | u_{k-1}, v_{k-1}, \nu^{k-1}, \rho^{k-1}) d(\nu_k, \rho_k) \qquad (6.38)$$

Then $\tilde{u}^*_{k-1}$ is found by

$$\gamma^*_k = \min_{\tilde{u}^*_{k-1}} \gamma_k \qquad (6.39)$$

where we assume $p(\mu_{k-1} | u_{k-1}, v_{k-1}, \nu^{k-1}, \rho^{k-1})$ and $p(\nu_k, \rho_k | u_{k-1}, v_{k-1}, \nu^{k-1}, \rho^{k-1})$ is known. Thus, we can determine the optimal strategies once $p(\mu_k | \nu^k, \rho^k)$ and $p(\nu_k, \rho_k | \nu^{k-1}, \rho^{k-1})$ are known. We compute these conditional densities in section 6.6. Again if there exist stationary strategies, the which are also equalibur strategies, then by theorem 6.4 we have the min max = max min condition at each stage.

$$\gamma_k^* = \min_{b^s_{u_{k-1}}} \max_{b^s_{v_{k-1}}} \gamma_k = \max_{b^s_{v_{k-1}}} \min_{b^s_{u_{k-1}}} \gamma_k \tag{6.40}$$

This equation will then be the analog of the Hamilton-Jacobi-Bellman partial differential equation for the Value function.

6.6 DERIVATION OF THE RECURSIVE RELATIONS FOR THE CONDITIONAL PROBABILITY DENSITY

Since the process $\{\nu_k\}$ as known by player I alone does not constitute a Markov Process, the derivation of the conditional probability density of the parts of a multi-dimensional Markov chain when the players observe only a particular process, the conditional probability densities can be derived based on the knowledge of $p(\chi_k|\chi_{k-1})$. However, we derive the conditional probability densities of μ_k or ρ_k based only on $\{\nu_k\}$.

Let $p_o(\chi_o) = p_o(\mu_o, \rho_o, \nu_o)$ be the a priori density. Since we require the conditioning to be based only on ν^k, while the densities depend on ν^k, ρ^k, we shall derive the two below. We have

$$p(\mu^k, \rho^k|\nu^k) = \frac{p(\chi^k)}{\int p(\chi^k)\, d\mu^k\, d\rho^k} \tag{6.41}$$

assuming a nonzero denominator and

$$p(\mu^k|\nu^k, \rho^k) = \frac{p(\mu^k, \rho^k|\nu^k)}{\int p(\mu^k, \rho^k|\nu^k)\, d\mu^k} \tag{6.42}$$

Hence given one form of density, we can go to the other density.

Now consider

$$p(\mu_k, \rho_k, \mu_{k+1}, \rho_{k+1}, \nu_{k+1} | \nu^k)$$

$$= p(\mu_k, \rho_k | \nu^k)\, p(\mu_{k+1}, \rho_{k+1}, \nu_{k+1} | \mu_k, \rho_k, \nu^k) \tag{6.43}$$

$$= p(\mu_k, \rho_k | \nu^k)\, p(x_{k+1} | x_k) \tag{6.43}$$

Eq. (6.43) follows from the Markovian property of the transitional density. Integrating both sides with respect to μ_k, ρ_k we have

$$p(\mu_{k+1}, \rho_{k+1}, \nu_{k+1} | \nu^k)$$

$$= p(\nu_{k+1} | \nu^k)\, p(\mu_{k+1}, \rho_{k+1} | \nu^k)$$

$$= \int p(\mu_k, \rho_k | \nu^k)\, p(x_{k+1} | x_k)\, d(\mu_k, \rho_k) \tag{6.44}$$

Hence

$$p(\mu_{k+1}, \rho_{k+1} | \nu^{k+1}) = \frac{\int p(\mu_k, \rho_k | \nu^k)\, p(x_{k+1} | x_k)\, d(\mu_k, \rho_k)}{\int \begin{bmatrix}\text{the above}\\ \text{numerator}\end{bmatrix} \times d(\mu_{k+1}, \rho_{k+1})} \tag{6.45}$$

Again consider

$$p(\nu_{k+1}, \rho_{k+1} | \nu^k, \rho^k)$$

$$= \int p(\mu_{k+1}, \nu_{k+1}, \rho_{k+1} | \nu^k, \rho^k)\, d\mu_{k+1}$$

$$= \int p(\mu_{k+1}, \nu_{k+1}, \rho_{k+1} | \mu^k, \nu^k, \rho^k)\, p(\mu^k | \nu^k, \rho^k)\, d\mu^{k+1}$$

$$= \int p(x_{k+1} | x_k)\, p(\mu^k | \nu^k, \rho^k)\, d\mu^{k+1} \tag{6.46}$$

Hence

$$p(\vartheta_{k+1}, \rho_{k+1}|\vartheta^k, \rho^k) = \int p(\chi_{k+1}|\chi_k) \frac{p(\mu^k, \rho^k|\vartheta^k)}{\int p(\mu^k, \rho^k|\vartheta^k) d\mu^k} d\mu^{k+1} \tag{6.47}$$

Equations (6.45), (6.47) define recursive relations governing the conditional probability densities with

$$p(\mu^0, \rho^0|\vartheta^0) = \frac{p(\chi_0)}{\int p(\chi_0) d\mu^0 d\rho^0} \tag{6.48}$$

In any particular application, these by themselves are unnecessary, once the form for the distribution function is assumed, since we compute instead recursive relations governing the sufficient statistics.

6.7 EXAMPLE

Consider the simple game

$$\begin{aligned} x_{k+1} &= ax_k + bu_k + cv_k + \eta_k \\ y_k &= x_k + \xi_k \\ z_k &= x_k + \zeta_k \end{aligned} \tag{6.49}$$

where x_k, u_k, v_k, y_k, z_k are all scalars, η_k, ξ_k, ζ_k are noise processes, a, b, c are known constants. The constraints on u, v are

$$\begin{aligned} U &= \{u : |u| \leq 1\} \\ V &= \{v : 0 \leq v \leq 1\} \end{aligned} \tag{6.50}$$

The game is assumed here to be a zero-sum two-person game. The payoff function from player I to player II is given by x_N^2 at

the end of the prespecified M stages.

$$I(u,v) = x_M^2 \tag{6.51}$$

Player I seeks to minimise the expected value while player II seeks to maximise the expected value. We consider two cases here.

(a) <u>The Game with No Noise</u> : Here set $\eta_k = \xi_k = \zeta_k = 0.$, That a mixed strategy solution is necessary at least for one of the players would be obvious in a single stage game where the payoff function is

$$\min_{u \in U} \max_{v \in V} x_1^2 = \min_{u \in U} \max_{v \in V} (ax_o + bu_o + cv_o)^2 \tag{6.52}$$

This is a particular case of convex games |70|, and hence the theorem given in the Appendix E holds. We find

$$\tilde{u}^* = u^o = - \text{sat}\,(ax_o + c/2) \tag{6.53}$$

$$\begin{aligned} \tilde{v}^* &= 1 \quad \text{if} \quad ax_o + bu_o < 1/2 \\ &= 0 \quad \text{if} \quad ax_o + bu_o > 1/2 \end{aligned} \tag{6.54}$$

Since the player does not know the exact value of u_1 , he chooses a mixed strategy:

$$\begin{aligned} \tilde{v}^* &= 1 \quad \text{with probability} \quad 1/2 \\ \tilde{v}^* &= 0 \quad \text{with probability} \ 1/2 \end{aligned} \tag{6.55}$$

For the multistage case the same is still a true, since there is no cost (payoff) attached to the intervening stages except v^* is now the semirandom binary jamming strategy with parameter 1/2.

The value of the game is $c^2/4$. Thus player II chooses the above stationary strategy while player I chooses a pure strategy.

(b) <u>The Game with Noise</u> : We now consider the system with known noise. Here,

$$E(\eta_k) = 0 = E(\xi_k), \quad \zeta_k = 0, \quad E(\eta_k^2) = q_k^2, \quad E(\xi_k^2) = r_k^2 \tag{6.56}$$

Consider the situation where player II chooses a semirandom binary stationary strategy first. Let

$$\text{Prob. } (v_k = 1) = \alpha \qquad \text{Prob. } (v_k = 0) = 1-\alpha \tag{6.57}$$

We determine the last stage strategy. Let

$$\mu_k = E\,(x_k|y^k) \qquad \Delta_k = \mathrm{Var}\,(x_k|y^k) \tag{6.58}$$

The last stage value γ_N is obtained thus:

$$\begin{aligned}
\gamma_N &= E\,(\,x_N^2 \mid y^{N-1}\,) \\
&= \int x_N^2\, p(x_N| x_{N-1}, u_{N-1}, v_{N-1}, a, b, c, y^{N-1}, \eta_{N-1}, \xi_{N-1}) \\
&\qquad \times\; p(x_{N-1}, a, b, c, \eta_{N-1}, \xi_{N-1} | y^{N-1})\, p(u_{N-1}|v_{N-1}, y^{N-1}) \\
&\qquad \times\; p(v_{N-1})\, d(x_N, x_{N-1}, a, b, c, \eta_{N-1}, \xi_{N-1}, u_{N-1}, v_{N-1}) \\
&= \int (a x_{N-1} + b u_{N-1} + c v_{N-1} + \eta_{N-1})^2\, p(x_{N-1}|y^{N-1})\, p(\eta_{N-1})\, p(\xi_{N-1}) \\
&\qquad \times\; p(u_{N-1}|v_{N-1}, y^{N-1})\, p(v_{N-1})\, d(x_{N-1}, u_{N-1}, v_{N-1}, \eta_{N-1}, \xi_{N-1})
\end{aligned}$$

$$= \int ((a\mu_{N-1} + bu_{N-1} + cv_{N-1})^2 + q^2_{N-1} + \Delta^2_{N-1}\, a^2)$$

$$\times\; p(u_{N-1}|v_{N-1}, y^{N-1})\; p(v_{N-1})\; d(u_{N-1}, v_{N-1})$$

$$= (a\mu_{N-1} + bu_{N-1})^2\,(1-\alpha) + (a\mu_{N-1} + bu_{N-1} + c)^2\,\alpha$$

$$+\; q^2_{N-1} + {}^{2}_{N-1}\, a^2$$

$$= (a\mu_{N-1} + bu_{N-1})^2 + 2c\alpha(a\mu_{N-1} + bu_{N-1})^2 + c^2\alpha$$

$$+\; q^2_{N-1} + \Delta^2_{N-1}\, a^2 \tag{6.59}$$

Since player I has a pure strategy, we have

$$u^o_{N-1} = -(c\alpha + a\mu_{N-1})/b \tag{6.60}$$

or if saturation is present

$$u^o_{N-1} = -\,\mathrm{sat}((c\alpha + a\mu_{N-1})/b) \tag{6.61}$$

Correspondingly the payoff can be written as

$$\gamma_N^* = c^2\alpha^2 - 2c^2\alpha^2 + c^2\alpha + q^2_{N-1} + \Delta^2_{N-1}\, a^2$$

$$= c^2(\alpha - \alpha^2) + q^2_{N-1} + a^2\Delta^2_{N-1} \tag{6.62}$$

which is a maximum if $\alpha = 1/4$. Hence

$$\gamma_N^* = c^2/4 + q^2_{N-1} + a^2\Delta^2_{N-1} \tag{6.63}$$

When u is saturated we have to maximise the expression

$$(a\mu_{N-1} - b \operatorname{sat}((c\alpha + a\mu_{N-1})/b)^2 + q^2_{N-1} + a^2 \Delta^2_{N-1} + c^2\alpha^2$$

$$+ 2\alpha\, (a\mu_{N-1} - b \operatorname{sat}((c\alpha + a\mu_{N-1})/b)^2 \qquad (6.64)$$

in α for player II.

Since in this problem by theorem 6.4 we can invert the minmax procedure we have

$$v_{N-1} = 1 \quad \text{if} \quad -((a\mu_{N-1} + bu_{N-1})/c) \le 1/2$$
$$= 0 \quad \text{if} \quad -((a\mu_{N-1} + bu_{N-1})/c) > 1/2 \qquad (6.65)$$

Let us choose $p(v_{N-1} = 1) = \alpha$. Then (6.59) is a maximum if $\alpha = 1/2$

In the general case we have

$$u^*_k = -\operatorname{sat}((a\mu_k + c/2)/b) \qquad (6.66)$$

$$p(v^*_k = 1) = 1/2 \quad , \quad p(v^*_k = 0) = 1/2 \qquad (6.67)$$

In the unconstrained case for u the Value can be computed as

$$J^*_k = c^2/4 + \Sigma^N_{i=k}\, q^2_{i-1} + \Sigma^N_{i=k}\, \Delta^2_{i-1}\, a^2 \qquad (6.68)$$

where Δ_i can be determined from (6.58).

In some combat situations the recovery process of disabled trucks, tanks, etc. in the presence of enemy fire is a problem for which the above example serves as a simplified model. The control v is the intermittent enemy fire. The potentiality x of units to be recovered is large and hence the

loss in unrecovered units is sought to be minimised at the end of N stages. This simplified model can also be interpreted as an electronic counter-measure problem.

6.8 CONCLUSIONS

The Markov Positional Game enabled us to identify different classes of mixed strategies under which the saddlepoint theorem is examined. The decomposition of the state space into four disjoint classes of states observed, inferred, remembered and ignored also brings out the difference between 'position' and 'state'. With incomplete structural information more concepts are needed which are given in the next chapter. Thus, the example considered cannot be solved for a saddlepoint condition under incomplete structural information in general. More appropriately by suitable modifications each player has an individual payoff. Hence a Nash equilibrium point is more appropriate.

VII N-PERSON MARKOV POSITIONAL GAMES

This chapter presents a generalisation of the results obtained in the last chapter and constitutes an attempt to partially answer the incomplete information problem raised in chapter 5. Thus in the N-player case, each player has different information sets and possibly payoff functions linked with some common constraints having stochastic elements. Since Kuhn's concept of an extensive game is valid for the N-person case, the extension is still possible within this framework. Hence the majority of results derived from the viewpoint of player I carry over to the results derived from the viewpoint of player i. Our first section below describes the transformation of the N-person incomplete structural information case into a game with complete structural information but incomplete position information, and perhaps imperfect information. Section 7.3 describes the process of obtaining the state vector for the N-person positional game with complete structural information, and this enables the construction of the N-person Markov Positional Game. Section 7.4 generalises the notion of a playable pair in deterministic differential games. In the last section the properties of various types of strategies are examined in terms of the new notions introduced in the previous section.

7.2 TRANSFORMATION OF THE INCOMPLETE INFORMATION GAME

We first consider an N-person positional game with

incomplete structural information. The game description as viewed by a player can be divided into three distinct classes:

(i) The umpire specifies to him the exact optimal strategy of some players and the exact nature of some of the system structural parameters.

(ii) The umpire specifies to him the optimal strategy of some other players of the set $\underline{N} = \{1, 2, \ldots, N\}$ in a probabilistic sense as also some other of the system structural parameters.

(iii) The umpire leaves him ignorant (does not specify) the strategies of the remaining players and the remaining structural parameters.

Let $\underline{U} = U^1, U^2, \ldots, U^N$, and

$\underline{u}^{c^i} \epsilon \underline{U}^{c^i}$ denote those collective strategies of some players known exactly to player i .

$\underline{u}^{p^i} \epsilon \underline{U}^{p^i}$ denote those collective strategies of some other players, known probabilistically to player i .

$\underline{u}^{u^i} \epsilon \underline{U}^{u^i}$ denote those strategies of the remaining players of which player i has complete uncertainty.

Obviously we must have

$$\underline{U}^{c^i} \times \underline{U}^{p^i} \times \underline{U}^{u^i} \times U^i = U^1 \times U^2 \times \ldots \times U^N \qquad (7.1)$$

Similarly let

$w^{e^i} \in W^{e^i}$ denote those parameters known exactly to player i.

$w^{f^i} \in W^{f^i}$ denote those parameters known probabilistically to player i.

$w^{q^i} \in W^{q^i}$ denote those parameters completely unknown to player i.

Then

$$W = W^{e^i} \times W^{f^i} \times W^{q^i} \qquad (7.2)$$

The position and observation equation of the system, as seen by player i, can be thus written as

$$\begin{aligned} x_{k+1} &= g_k(x_k, u_k^i, \underline{u}_k^{e^i}, \underline{u}_k^{p^i}, \underline{u}_k^{u^i}, w_k^{e^i}, w_k^{f^i}, w_k^{q^i}) \\ y_k^i &= h_i(x_k, w_k^{e^i}, w_k^{f^i}, w_k^{q^i}) \end{aligned} \qquad (7.3)$$

Let player i consider a subjective probability description for $\underline{u}^u$, w^q and split his resources into u^{i1} and u^{i2} such that $u^i \in U^i$ and

$$\begin{aligned} u^i &= \beta_1^i u^{i1} + \beta_2^i u^{i2} \\ u^{i1} &\in U^i, \; u^{i2} \in U^i \end{aligned} \qquad (7.4)$$

Let u^{i1} be utilised such that the specified objective function is maximised with a supercriterion, while u^{i2} be utilised to minimise a subjective payoff function. Let us consider these to be in the hands of two 'Agents' of player i. As each agent then views the game he has only a single payoff function and elements either completely known or with specified

probability distributions, i.e., to each agent the game appears now as one with certainty or risk. The corresponding theory of deterministic or Markov Positional Games can be applied in principle to determine the optimal strategies of the agents. We have thus shown heuristically the existence of a game with certainty or risk, but incomplete position information as equivalent to a game with certainty, risk and uncertainty and which has incomplete structural information. We can, thus, study the properties of the former game. We can conclude the above discussion in the form of a theorem.

<u>Theorem 7.1</u> : Given an incomplete structural information positional game it is possible to find an equivalent incomplete position information positional game such that the given payoff function of the first game is replaced by an equivalent payoff function with a supercriterion, together with a subjective payoff function for each player which is arbitrary and the optimal strategy for which has no effect on the dynamics of the position of the game. In the limit, when the incomplete information game tends to a complete information game, the corresponding payoff functions coincide at the optimal point and the strategies with supercriterion tend to the strategies of the complete structural information game.

<u>Corollary 7.1</u> : If the incomplete structural information game has complete position information then the above limiting process leads to a game with complete information.

7.3 STATE DESCRIPTION

We next consider the problem of determining the state vector for player i in an N-person game with complete structural information. As the ith player views the game, the description appears to him to be

$$x_{k+1} = g_k(x_k, u_k^i, r_k^i, \xi_k^i)$$
$$y_k^i = H_k^i(x_k, \eta_k^i) \tag{7.5}$$

Let

$$u^{ik} \triangleq (u_1^i, u_2^i, \ldots, u_k^i)$$
$$r^{ik} \triangleq (r_1^i, r_2^i, \ldots, r_k^i)$$
$$y^{ik} \triangleq (y_1^i, y_2^i, \ldots, y_k^i)$$
$$\xi^{ik} \triangleq (\xi_1^i, \xi_2^i, \ldots, \xi_k^i) \tag{7.6}$$
$$\eta^{ik} \triangleq (\eta_1^i, \eta_2^i, \ldots, \eta_k^i)$$
$$\zeta^{ik} \triangleq (\zeta_1^i, \zeta_2^i, \ldots, \zeta_k^i)$$

and the payoff to the i-th player at the end of N stages is given by

$$I^i(\underline{u}) = E\left\{ \Sigma_{k=1}^{N} L_k^i(x_k, u_k^i, r_k^i, \xi_k^i, \eta_k^i, \zeta_k^i) \right\} \tag{7.7}$$

and his objective is to maximize $I^i(\underline{u})$ for $\{u_k^i \epsilon U_k^i\}$. Then the state vector for this player can be considered to be

$$\phi_k^i = (x_k, u^{ik}, r^{ik}, \xi^{ik}, \eta^{ik}, y^{ik}, \zeta^{ik}) \tag{7.8}$$

Let a reduced state vector be found as

$$\bar{\phi}_k^i = (x_k, u_k^i, s_k^i, y_k^i) \tag{7.9}$$

where the state vector s_k^i is obtained from a recursive equation

$$s_{k+1} = f_k(y_{k+1}, s_k, u_k) \tag{7.10}$$

such that

$$\text{Prob.}\ (\ \phi_k^i\ ) = \text{Prob.}\ (\ \bar{\phi}_k^i\ ) \tag{7.11}$$

and

$$\text{Prob.}\ (\ \phi_k^i \mid \phi^{i(k-1)}) = \text{Prob.}\ (\ \phi_k^i \mid \phi_{k-1}^i\ )$$

$$= \text{Prob.}\ (\ \bar{\phi}_k^i \mid \bar{\phi}^{i(k-1)}) = \text{Prob.}\ (\ \bar{\phi}_k^i \mid \bar{\phi}_{k-1}^i\ ) \tag{7.12}$$

As viewed by player i, the state vector can be decomposed for him into four classes given in theorem 6.1. The state for the umpire is then

$$\chi_k = \bigcup_{i \in \underline{N}} (\bar{\phi}_k^i) \tag{7.13}$$

where $\underline{N} = \{1, 2, \ldots, N\}$ (7.14)

In this manner, we arrive at the N-person Markov Positional Game

$$\begin{aligned} \chi_{k+1} &= f_k (\chi_k, \underline{u}_k) \\ y_k^i &= \rho_k^i \end{aligned} \tag{7.15}$$

Let $x_k \in \mu_k$ then

$$\mu_k \cup \left(\bigcup_{i \in \underline{N}} \rho_k^i \right) = x_k \tag{7.16}$$

The payoff function of the ith player can then be written as

$$I^i(\underline{u}_k) = \Sigma_{k=1}^{M} L^i(x_k, \underline{u}_k^i) \tag{7.17}$$

We then have the theorem

Theorem 7.2 : <u>A decomposition of the state space as viewed by the ith player of an N-person Markov Positional Game into the subclasses observed, inferred, remembered and ignored can be made in general.</u>

7.4 PLAYABILITY

Just as the notion of controllability in one-sided control problems plays an important role, the notion of playability is important in deriving the maximal class of strategies for all players with which termination of the game in the prescribed sense can be assured. The concept of a playable pair introduced by Berkovitz |71| is the prime motivation for the generalizations presented here. In the game considered by Berkovitz, given the sets U, V of the admissible strategies of the two players, $u \in U$ is said to be a playable strategy for player I if the pair (u,v) assures termination of the game starting at $x(o)$ for all $v \in V_o$ V. Similarly $v \in V$ is said to be a playable strategy for player II if the player (u,v) assures termination of the

game starting at $x(0)$ for all $u \in U_o$. The pair (U_o, V_o) then constitutes the pair of the sets of all <u>mutually</u> playable strategies. In this, it is assumed that the sets U and V are completely specified to both players.

In the general case, we first need an understanding of what we mean by the termination of the game.

<u>Definition 7.1</u> :

<u>Termination of the Game</u>: As viewed by a player i if through the observations granted to him he can assure transfer of the position of the system starting at $x(0) = x_0$ to the terminal position x_N, $N < \infty$, for some fixed strategies of the remaining players, then we say that the termination of the game in the prescribed sense can be assured.

Let us now consider the viewpoint of player i when the game specified to him is of incomplete structural information, before <u>he hands over the game to his agents.</u> Here he should be assured of termination of the game in the prescribed sense if the other players care to choose different strategies. In section 7.2 we have shown that as viewed by player i there exists a decomposition of $\underset{i \in \underline{N}}{\times} U^i$ -space into $\underline{U}^{o^i}$, $\underline{U}^{p^i}$, $\underline{U}^{u^i}$, U^i. However, the player need not have complete knowledge of $\underline{U}^{p^i}$, $\underline{U}^{u^i}$.

Let us now find the mappings which determine the (subjective and probabilistic) knowledge of player i of the sets $\underline{U}^{o^i}$, $\underline{U}^{p^i}$, $\underline{U}^{u^i}$. We first define

$$\underline{P}^i = \Sigma_k r^k \quad k\in\underline{P}$$

$$\underline{M}^i = \Sigma_k r^k \quad k\in\underline{M} \tag{7.18}$$

$$\underline{I}^i = \Sigma_k r^k \quad k\in\underline{C}$$

where

$$\underline{P} \cup \underline{M} \cup \underline{C} \cup i = \underline{N} \tag{7.19}$$

and $\underline{P}$, $\underline{M}$, $\underline{C}, i$ are pairwise disjoint. Let

$$\underline{U}^{p^i} \subseteq \mathbb{R}^{P^i}, \quad \underline{U}^{u^i} \subseteq \mathbb{R}^{M^i}, \quad \underline{U}^{c^i} \subseteq \mathbb{R}^{I^i} \tag{7.20}$$

Let us define the mappings

$$c : \underline{U}^{c^i} \to \mathcal{C}^i \subseteq \mathbb{R}^{I^i}$$

$$p : \Omega^{p^i} \times \underline{U}^{p^i} \to \mathcal{P}^i \subseteq \mathbb{R}^{P^i} \tag{7.21}$$

$$m : \Omega^{u^i} \times \underline{U}^{u^i} \to \mathcal{M}^i \subseteq \mathbb{R}^{M^i}$$

where the Ω^{p^i} space is the set of all randomizations over $\underline{U}^{p^i}$ known to player i, Ω^{u^i} is the set of all randomizations over $\underline{U}^{u^i}$ subjectively formed by player i. The mapping c^i corresponds to the certainty map of the i-th player's viewpoint; p^i is the 'risk' map, i.e., the mapping by which player i knows about $\underline{U}^{p^i}$ with known probabilities; m^i is the map by which player i subjectively constructs the set $\mathcal{M}^i$ to determine $\underline{U}^{u^i}$.

<u>Definition 7.2</u>: As viewed by player i strategy $u^i \in U^i$ is playable definitely against the strategy $u^j \in U^j$ for some $j\in\underline{C}$,

if the pair $(u^i, \ell^i(u^j))$, where the image of u^j under the map ℓ^i is $\ell^i(u^j)$, assures termination of the game in the prescribed sense when the strategies u^k $k \neq i,j$ assure collectively the termination of the game in the prescribed sense for some fixed u^k $k \neq i,j$, for all possible $\ell^i(u^j) \in \mathcal{L}_0^i \subseteq \mathcal{L}^i$.

<u>Definition 7.3</u>: As viewed by player i, strategy $u^i \in U^i$ is playable with probability p_j^i against the strategy $u^j \in U^j$ for some $j \in \underline{K}$ if the pair $(u^i, {}^i(u^j))$, where the image of u^j under the map p^i is $p^i(u^j)$, assures termination of the game in the prescribed sense with probability p_j^i when the strategies u^k $k \neq i,j$ assure collectively the termination of the game in the prescribed sense for some fixed u^k, $k \neq i,j$, for all possible $p^i(u^j) \in \mathcal{P}_0^i \subseteq \mathcal{P}^i$.

<u>Definition 7.4</u>: As viewed by player i, strategy $u^i \in U^i$ is playable subjectively against $u^j \in U^j$ for some $j \in \underline{M}$ if there exists a finite subjective nonzero probability μ_j^i such that the pair $(u^i, m^i(u^j))$, where the image of u^j under the map m^i is $m^i(u^j)$, assures termination of the game in the prescribed sense with probability μ_j^i, when the strategies u^k $k \neq i,j$ assure termination of the game in the prescribed sense for some fixed strategies u^k $k \neq i,j$ for all $m^i(u^j) \in \mathcal{M}_0^i \subseteq \mathcal{M}^i$, $j \in \underline{M}$.

<u>Definition 7.5</u>: As viewed by player i, strategy $u^i \in U^i$ is

playable provided there exist sets $\mathcal{M}_o^i \subseteq \mathcal{M}^i, \mathcal{P}_o^i \subseteq \mathcal{P}^i, \mathcal{C}_o^i \subseteq \mathcal{C}^i$ such that u^i is playable against any strategy $m^i(u^{u^j}) \in \mathcal{M}_o^i$, $j \in \underline{M}$, with non-zero subjective probabilities $\mu_j^i(u^j)$; $p^i(u^{p^j}) \in \mathcal{P}_o^i$, $j \in \underline{P}$ with non-zero probabilities $p_j^i(u^j)$; $c^i(u^{c^j}) \in \mathcal{C}_o^i$, $j \in \underline{C}$ with certainty; in the sense that the tuple $(u^i, \underline{u}^{c^i}, \underline{u}^{p^i}, \underline{u}^{u^i})$ assure termination of the game in the prescribed sense.

Let

$$
\begin{aligned}
c^{-1}(\mathcal{C}_o^i) &= \mathcal{C}_o^{i^{-1}} \\
p^{-1}(\mathcal{P}_o^i) &= \mathcal{P}_o^{i^{-1}} \qquad (7.22) \\
m^{-1}(\mathcal{M}_o^i) &= \mathcal{M}_o^{i^{-1}}
\end{aligned}
$$

For some $j \in \underline{C}^i$, $j \neq i$, one can find some strategy $u^j \in U^j$ playable against u^i from the viewpoint of player i. Corresponding to this, from the viewpoint of player j if u^i is playable either subjectively or objectively or with certainty with the strategy u^j we try to determine an equivalent set $\mathcal{C}_o^j$ such that $c^j(u^i) \in \mathcal{C}_o^j$ and $(u^j, c^j(u^i))$ is a playable pair from the viewpoint of player i. The notation $\mathcal{B}_o^j$ stands for the subjective, risk or certainty map.

<u>Definition 7.6:</u> Two strategies $u^i \in U^i$, $u^j \in U^j$ are said to be mutually playable with certainty if there exists certainty maps $\mathcal{C}_j^i: U^j \to \mathcal{R}^{r^j}$ and $\mathcal{C}_i^j: U^i \to \mathcal{R}^{r^i}$ such that

$u^i \in \ell^j_{i_0} \subseteq \ell^j_i$ and $u^j \in \ell^i_{j_0} \subseteq \ell^i_j$ are arbitrary and such that the remaining strategies of players $k \neq i,j$ assure termination of the game in the prescribed sense from the viewpoint of both players.

__Definition 7.7:__ Two strategies $u^i \in U^i$, $u^j \in U^j$ are said to be mutually playable with 'risk' probabilities $(p^j_i, p^i_j) > 0$ if there exist risk maps $\rho^i_j : U^j \times \Omega^j \to \mathbb{R}^{r^j}$, $\rho^j_i : U^i \times \Omega^i \to \mathbb{R}^{r^i}$ such that $u^i \in \mathcal{P}^j_{i_0} \subseteq \mathcal{P}^j_i$ and $u^j \in \mathcal{P}^i_{j_0} \subseteq \mathcal{P}^i_j$ are arbitrary and such that the remaining strategies assure collectively the termination of the game in the prescribed sense from the viewpoint of both players.

__Definition 7.8:__ Two strategies $u^i \in U^i$, $u^j \in U^j$ are said to be mutually playable with subjective probabilities $(\mu^i_{j0}, \mu^j_{i0}) > 0$ if there exist subjective maps $m^j_i : U^i \times \Omega^i \to \mathbb{R}^{r^i}$, $m^i_j : U^j \times \Omega^j \to \mathbb{R}^{r^j}$ such that $u^i \in \mathcal{M}^j_{i_0} \subseteq \mathcal{M}^j_i$ and $u^j \in \mathcal{M}^i_{j_0} \subseteq \mathcal{M}^i_j$ are arbitrary and such that the remaining strategies assure collectively the termination of the game in the prescribed sense from the viewpoint of both players.

__Theorem 7.3:__ A mutually playable pair of strategies exist if and only if $u^i \in \mathcal{B}^j_i \subseteq \mathcal{B}^j_i \Leftrightarrow u^j$ $\mathcal{B}^i_{j_0} \subseteq \mathcal{B}^i_j$ where $\mathcal{B}^i_j : U^j \to \mathbb{R}^{r^j}$, $\mathcal{B}^j_i :$ $U^i \to \mathbb{R}^{r^i}$ and $\mathcal{B}$ is either a certainty, risk or subjective map.

Proof: Suppose $u^i \in \mathcal{B}^i_{j_o}$ is mutually playable with $u^j \Leftrightarrow$ All $u^i \in \mathcal{B}^j_{i_o}$ are playable with $u^j \in \mathcal{B}^i_{j_o} \Leftrightarrow$ All $u^j \in \mathcal{B}^i_{j_o}$ are playable with $u^i \in \mathcal{B}^j_{i_o} \Leftrightarrow u^j \in \mathcal{B}^i_{j_o}$ is a mutually playable pair with $u^i \in \mathcal{B}^j_{i_o}$. Q.E.D.

Corollary to Theorem 7.3: If $\mathcal{B}^i_j = \mathcal{M}^i_j$, $\mathcal{B}^j_i = \mathcal{M}^j_i$, then the theorem is true for the subjectively mutually playable pair, with subjective probabilities $(\mu_i, \mu_j) > 0$.

If $\mathcal{B}^i_j = \mathcal{P}^i_j$, $\mathcal{B}^j_i = \mathcal{P}^j_i$, then the theorem is true for the mutually playable pair with objective probabilities $(p_i, p_j) > 0$.

If $\mathcal{B}^i_j = \mathcal{C}^i_j$, $\mathcal{B}^j_i = \mathcal{C}^j_i$, then the theorem is true for the mutually playable pair with certainty.

Definition 7.9: If strategy $u^i \in U^i$ is playable for all i such that $u^i \in \mathcal{B}^i_j$, $j \neq i$ $\mathcal{B}^i_j$ is either the certainty, risk or subjective map, then $\underline{u}$ is completely mutually playable.

Theorem 7.4: A minimal class of strategies $(U^1_o, \ldots, U^N_o)$ exist such that viewed with certainty or risk or subjectively, a member $u^i \in U^i_o$ is playable against all the others' in the sense of definition 7.5, for every $i = 1, \ldots, N$.

Proof: Let $u^i \in U^i$ then it is possible to find certainty, risk and subjective maps for player i to view the sets U^j $j \neq i$. If it is mutually playable then we have a set $U^i_j \subseteq U^i$ such that $u^i \in U^i_j$ is mutually playable with $u^j \in U^j_i \subseteq U^j$ in the sense of definitions 7.6 – 7.8. This

determines a set U_j^i for player i considered with respect to player j. Let

$$U_o^i = \;\; {}_{j\in N,\, j\neq i} \;\; U_j^i \tag{7.23}$$

Then $u^i \in U_o^i$ is playable in the sense of definition 7.5. Since we had i arbitrary, this is true for all $i \in N$. Hence the collection $(U_o^1,\ldots,U_o^N)$ is minimal. Q.E.D.

Theorem 7.8: In a game with complete structural information only playability of strategies in the sense of certainty or risk are relevant.

Proof: This follows, since in a complete structural information game the players either know with certainty the various strategies or with risk. If there exist stochastic elements in the system and observation equations, these have objective probabilities. If any uncertainty in the complete sense exists, it has been replaced by the objective probabilities of an agent (equivalent to the subjective probabilities of the player) so that no subjective maps are necessary. Q.E.D.

With these definitions many more results can possibly be derived on playability in N-person games. For instance, much more structure can be built into the game description by bringing in topological considerations which have been avoided in this thesis.

7.6 PROPERTIES OF DIFFERENT TYPES OF STRATEGIES

Just as in the two-person zero-sum complete structural

information case we could describe under certain conditions the notions of mixed, behaviour, stationary and pure strategies, we can as well describe these for the general N-person game. Since these are independent of the payoff functions and the concept of playability, the definitions carry over to the N-person game for these four cases. We thus need to examine only the notions of Bayes, extended Bayes, inf-sup and equaliser strategies.

<u>Definition 7.10</u>: Let $I^i(\underline{u})$ denote the payoff to player i due to a tuple of strategies playable for player i. Then a strategy $u^{i*} \in U_0^i \; U^i$ is said to be a Bayes playable strategy for player i if it satisfies

$$I^i(\underline{u};u^{i*}) = \inf_{u^i \in U_0^i} I^i(\underline{u}) \tag{7.24}$$

If this is true for every $i \in \underline{N}$ then we say that $\underline{u}^*$ is an <u>equilibrium point playable strategy for each player from his individual viewpoint</u>. Clearly this is the Nash equilibrium point.

<u>Theorem 7.6</u>: If $\underline{u}^*$ is an equilibrium point tuple of strategies then $\underline{u}^*$ is a completely mutually playable tuple.

<u>Proof</u>: $\underline{u}^*$ is an equilibrium point tuple of strategies. $\Rightarrow$ For every i there exists a $u^{i*} \in U^i$ such that $I^i(\underline{u}) \geq I^i(\underline{u};u^{i*})$ $\Rightarrow$ u^{i*} is Bayes playable with respect to u^j and u^{j*} is Bayes playable with respect to u^i, $i,j \in \underline{N}$. $\Rightarrow$ (u^i,u^j) is a mutually playable pair for all $i,j \in \underline{N}$ $\Rightarrow \underline{u}^*$ is a completely mutually playable tuple. Q.E.D.

We have now the notion of an extended Bayes strategy for each player.

<u>Definition 7.11</u>: If for some $i \in \underline{N}$, and $\epsilon > 0$ we have

$$I^i(\underline{u}_g;\, u^{i*}) \leq \inf_{u^i \in U_0^i} I^i(\underline{u}_g;\, u^i) + \epsilon \tag{7.25}$$

u^{i*} is called an extended Bayes strategy for player i.

<u>Definition 7.12</u>: If there exists $u^{i*} \in U_0^i$ such that

$$I^i(\underline{u};\, u^{i*}) = C \tag{7.26}$$

where C is a fixed constant, for any playable strategy for player i then u^{i*} is an equaliser strategy for player i.

We have the following analogs of theorems 6.2 - 6.4.

<u>Theorem 7.7</u>: From the viewpoint of player i he has a playable stationary strategy u^{io} such that $u^{io} \in U_B^i$, the set of all behaviour strategies for player i we have

$$I^i(\underline{u};\, u^{io}) = \max_{u^{io} \in U_B^i} \hat{I}(\underline{u}) \tag{7.27}$$

<u>Theorem 7.8</u>: If $u^{io} \in U_o^i$ is an equaliser strategy, is also extended Bayes and is playable then it is also an equilibrium point strategy for player i. If it is true for every player $i \in \underline{N}$ then the game has a Nash equilibrium point.

<u>Theorem 7.9</u>: Let $\underline{u}^*$ constitute a completely mutually playable tuple of strategies. If $\underline{u}^*$ are also stationary strategies for each player and extended Bayes for each player, then the

tuple $\underline{u}^*$ are equilibrium point strategies and the N-person game has an equilirbium point in mixed strategies where only equaliser strategies need be used by each player.

Thus, for a Nash equilibrium point to hold in a general N-person game, we must have a completely mutually playable tuple of strategies which are equalisers and stationary. In the cooperative case the mutual playability will be interesting to investigate. For it will tell when a player should enter into cooperation when subjective factors play a role.

The determination of the optimal strategies of a player and the recursive equations for the conditional probability densities is no problem. The relations in sections 6.5 - 6.6 are carried over to the viewpoint of a general player i.

Within the framework thus evolved, it should be possible to find a solution to the generalised trainer-learner problem.

7.6 CONCLUSIONS

In this chapter we have outlined some of the conceptual problems in extending the Markov Positional Game theory to the N-person case. With assumptions of continuity of functions, convexity of payoff functions, etc. many more results are obtainable.

This chapter also concludes our studies in differential games in this thesis.

VIII CONCLUSIONS

In these studies, without resorting to topological arguments and many more strict mathematical techniques, we have been able to show the intimate relation between the positional game and the differential game. Different problem areas have been identified through the introduction of many concepts which were hitherto absent in differential game theory. The motivation for many problems has been derived from existing engineering problems. In particular, the conceptual framework developed in this thesis is suited to evolve a more comprehensive theory of differential games with incomplete information. Each chapter has many possible extensions and developments- all worthwhile investigating, and these have been pointed out at the appropriate places in the thesis.

Apart from these, other related problems are:

(i) the continuous Markov Positional Game and the nature of different types of strategies in this case,

(ii) the computational aspects of the positional and differential game problems,

(iii) the counterpart of singular solution problems of control theory,

(iv) the synthesis from open loop policies of control strategies, in the presence of observation constraints, noise, etc.,

(v) stability considerations of playable strategies,

(vi) playability in N-person linear differential games under cooperation and non-cooperation.

These have not been considered herein and should be quite challenging for the interested investigator.

LIST OF REFERENCES

1. Isaacs, R., Differential Games (book), New York, John Wiley and Sons, 1965.

2. Berkovitz, L.D., 'A Survey of Differential Games' in Mathematical Theory of Control (book), A.V. Balakrishnan and L.W. Neustadt (Eds.), New York, Academic Press, 1967.

3. Pontryagin, L.S., 'Linear Differential Games' in Mathematical Theory of Control (book), A.V. Balakrishnan and L.W. Neustadt (Eds.), New York, Academic Press, 1967.

4. Shuford, E.H., 'Some Bayesian Learning Processes' in Human Judgements and Optimality (book), M.W. Sholley II and G.L. Bryan (Eds.), New York, John Wiley and Sons, 1964.

5. Ragade, R.K., 'Game Theoretic Formulation of Optimal Control Problems', Master's Thesis, Department of Electrical Engineering, Indian Institute of Technology, Kanpur.

6. Von Neumann, J. and Morgenstern, O., Theory of Games and Economic Behaviour (book), Princeton, Princeton University Press, 1946

7. Milnor, J., 'Sums of Positional Games' in Contributions to the Theory of Games Vol. II (book), H.W. Kuhn and A.W. Tucker (Eds.), Study No.28, Annals of Maths. Studies, Princeton, Princeton University Press, 1953.

8. Mycielski, J., 'Continuous Games with Perfect Information' in Advances in Game Theory (book), M. Dresher, L.S. Shapley and A.W. Tucker (Eds.), Study No.52, Annals of Maths. Studies, Princeton, Princeton University Press, 1964.

9. Kuhn, H.W., 'Extensive Games and the Problem of Information' in Contributions to the Theory of Games Vol. II (book), H.W. Kuhn and A.W. Tucker (Eds.), Study No.28, Annals of Maths. Studies, Princeton, Princeton University Press, 1953.

10. Dalkey, N., 'Equivalence of Information Patterns and Essentially Determinate Games' in Contributions to the Theory of Games Vol. II (book), H.W. Kuhn and A.W. Tucker (Eds.), Study No.28 , Annals of Maths. Studies, Princeton, Princeton University Press, 1953.

11. Davis, M., 'Infinite Games of Perfect Information' in <u>Advances in Game Theory</u> (book), M. Dresher, L.S. Shapley and A.W. Tucker (Eds.), Study No.52, Annals of Maths. Studies, Princeton, Princeton University Press, 1964.

12. Fleming, W.H., 'Optimal Control of Diffusion Processes' in <u>Functional Analysis and Optimization</u> (book), E. Caianiello (Ed.), New York, Academic Press, 1966.

13. Shapley, L.S., 'Stochastic Games', <u>Proc. Natl. Acad. Sci. U.S.A.</u>, <u>39</u>, 1953, pp 1095-1100.

14. Everett, H., 'Recursive Games' in <u>Contributions to the Theory of Games Vol. III</u> (book), M. Dresher, A.W. Tucker and P. Wolfe (Eds.), Study No.39, Annals of Maths. Studies, Princeton, Princeton University Press, 1957.

15. Zachrisson, L.E., 'Markov Games' in <u>Advances in Game Theory</u> (book), M. Dresher, L.S. Shapley and A.W. Tucker (Eds.), Study No.52, Annals of Maths. Studies, Princeton, Princeton University Press, 1964.

16. Howard, R.A., <u>Dynamic Programming and Markov Processes</u> (book), New York, John Wiley and Sons, 1960.

17. Petrosyan, L.A., 'Differential Games of Survival with Many Participants', <u>Dokl. Akad. Nauk. S.S.S.R.</u> <u>161</u> (1965), pp 285-287.

18. Karvovskiy, G.S. and Kuznetsov, A.D., 'A Maximum Principle for N-Players', <u>Engg. Cybernetics</u>, June 1966, pp 10-14.

19. Bollinger, R.L. et al. (Eds.), <u>Optimalizing and Adaptive Control</u> (book), Proceedings of the First International Symposium, Rome, Italy, 1962. Instrument Society of America, 1962.

20. Tsypkin, Ya. Z., 'Adaptation, Training and Self-Organization in Automatic Systems', <u>Automation and Remote Control</u> (English Translation) Vol. 27, No.1, Jan. 1966, pp 16-51.

21. Bellman, R.E., <u>Adaptive Control Processes: A Guided Tour</u> (book), Princeton, Princeton University Press, 1961.

22. Fu, K.S. and Gibson, J.E., 'Philosophy and the State of the Art of Learning Control Systems', <u>Tech. Report TR-EE-63-7</u> of the School of Electrical Engineering, Purdue University, Lafayette, Indiana, Nov. 1963.

23. Sworder, D.D., <u>Optimal Adaptive Control Systems</u> (book), New York, Academic Press, 1966.

24. Fel'dbaum, A.A., Optimal Control Systems (book), New York, Academic Press, 1965.

25. Witsenhausen, H.S., 'Minmax Control of Uncertain Systems', Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, Mass., 1966.

26. Bellman, R.E., 'Challenges of Modern Control Theory', Int. J. Engg. Sci. Vol. 4, 1964.

27. Karlin, S., Mathematical Methods in Games, Programming and Economics Vol. II (book), Reading, Addison Wesley, 1966.

28. Wang, P.K.C., 'Optimal Control of Distributed Parameters Systems' in Advances in Control Systems Vol. I (book), C.T. Leondes (Ed.), New York, Academic Press, 1964.

29. Balakrishnan, A.V., 'Semigroup Theory and Control Theory', Proceedings of the IFIP Congress, 1965.

30. Kursunoğlu, B., Modern Quantum Mechanics (book), San Fransisco, W.H. Freeman and Co., 1962.

31. Koenig, H.E., Tokad, Y., Kesavan, H.K., Analysis of Discrete Physical Systems (book), New York, McGraw-Hill, 1967

32. Harsanyi, J.C., 'Games with Incomplete Information Played by " Bayesian " Players I-III, Part I, The Basic Model,' Management Science Vol. 14, No.3, Nov. 1967, pp 159-182.

33. Sakaguchi, M., 'Information Pattern, Learning Structure and Optimal Decision Rule', J. Information and Control, Vol. 6, No.1, 1963, pp 218-224.

34. Kirillova, F.M., 'Applications of Functional Analysis to the Theory of Optimal Processes', J. SIAM CONTROL, Vol.5, No.1, 1967.

35. Balakrishnan, A.V., 'Operator Theoretic Formulation of a Class of Control Problems and Steepest Descent Method of Solution', J. SIAM CONTROL, Vol. 1, No.2, 1963.

36. Gindes, V.B., 'Optimal Control in a Linear System Under Conflict', J. SIAM CONTROL, Vol.5, No.1, 1967.

37. Porter, W.A., 'On Function-Space Pursuit Evasion Games', J. SIAM CONTROL, Vol.5, No.4, 1967.

38. Sakawa, Y., 'On Linear Differential Games', in Mathematical Theory of Control (book), A.V. Balakrishnan and L.W. Neustadt (Eds.), New York, Academic Press, 1967.

39. Appelgren, L., 'An Attrition Game', Operations Research, Vol.15, No.1, Jan-Feb. 1967.

40. Hsieh, H.C., 'Synthesis of Adaptive Control Systems by Function Space Methods' in Advances in Control Systems Vol. II (book), C.T. Leondes (Ed.), New York, Academic Press, 1965.

41. Ragade, R.K., 'Function Space Formulation of Differential Game Problems' in Selected Topics in Games and Decisions, I.G. Sarma and K.R. Sarma (Eds.), Research Report No.EE-67-1, Department of Electrical Engineering, Indian Institute of Technology, Kanpur, May 1967.

42. Hurwicz, L., 'Programming in Linear Spaces' in Studies in Linear and Non-linear Programming (book), K.J. Arrow, L. Hurwicz and H. Uzawa (Eds.), Stanford, Stanford University Press, 1958.

43. Lack, G.N.T., 'Optimization Studies with Applications to Planning in the Electric Power Industry and Optimal Control Theory', Doctoral Dissertation, Stanford University, 1965.

44. Kalman, R.E., 'On the General Theory of Control Systems' in Proceedings of the First IFAC, Moscow, 1960, London, Butterworths, 1962.

45. Sarma, I.G., Ragade, R.K., Prasad, U.R., 'Necessary Conditions for the Existence of Optimal Strategies in Non-Cooperative N-Person Differential Games' (To be published).

46. Sarma, I.G., Ragade, R.K., Prasad, U.R., 'N-Person Differential Games', to be presented at the Symposium on Control Theory, Indian Institute of Science, Bangalore, May 16-18, 1968.

47. Sarma, I.G., Ragade, R.K., Prasad, U.R., 'A Differential Game Model for the Zeuthen-Nash Bargaining Problem' (To be published).

48. Prasad, U.R., 'Differential Games with Incomplete Information', Forthcoming Ph.D. Thesis, Department of Electrical Engineering, Indian Institute of Technology, Kanpur.

49. Luce, R., and Raiffa, H., Games and Decisions (book), New York, John Wiley and Sons, 1957.

50. Berkovitz, L.D., 'Necessary Conditions for Optimal Strategies in a Class of Differential Games and Control Problems', J. SIAM CONTROL, Vol. 5, No.1, 1967.

51. Athans, M. and Falb, P.L., Optimal Control (book), New York, McGraw-Hill, 1966.

52. Nelson, W.L., 'On the Use of Optimization Theory for Practical Control System Design', IEEE Trans. on Automatic Control, Vol.AC-9, Oct. 1964, pp 469-477.

53. Hall, A.D., A Methodology for Systems Engineering (book), Princeton, D. Van Nostrand & Co., 1962.

54. Machol, R.E. (Ed.), Systems Engineering Handbook (book), New York, McGraw-Hill, 1965.

55. Hsu, J.C. and Meserve, W.E., 'Decision-Making in Adaptive Control Systems', IRE Trans. on Automatic Control, PGAC-7, Jan. 1962, pp 24-32.

56. Rosenfeld, J.L., 'Adaptive Decision Processes', Technical Report 403, Sept. 27, 1962, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Mass., U.S.A.

57. Ackoff, R.L., Scientific Method (book), New York, John Wiley and Sons, 1962.

58. Sakaguchi, M., 'Some Examples of Bayesian Adaptive Programming', J. Information and Control, Vol.9, 1966, pp 230-245.

59. Fishburn, P.C., 'Analysis of Decisions with Incomplete Knowledge of Probabilities', J. Operations Research, Vol.13, No.2, March-April 1965.

60. Harsanyi, J.C., 'A General Theory of Rational Behaviour in Game Situations', Econometrica, Vol.34, No.3, July 1966.

61. Murphy, R.E., Adaptive Economic Processes (book), New York, Academic Press, 1965.

62. Swingle, P.G., 'The Effect of the Win-Loss Difference upon Cooperative Responding in a Dangerous Game', J. Conflict Resolution, Vol.11, No.2, June 1967.

63. McLane, P.J. and Graeffe, P.W.U., 'Optimal Regulation of a Class of Linear Stochastic Systems with Quadratic Criteria', INT. J. CONTROL, Vol.5, No.2, 1967, pp 135-143.

64. Meier, Louis III, 'Combined Optimum Control and Estimations Theory', Tech. Report NAS2-2457, Stanford Research Institute, Menlo Park, California, Oct. 1966.

65. Sancho, N.G.F., 'Optimization of Linear Stochastic Control Systems Operating Over a Finite Time Interval', INT. J. CONTROL, Vol.3, No.4, 1966, pp 313-334.

66. Aumann, R., 'Mixed and Behaviour Strategies in Infinite Extensive Games', in Advances in Game Theory (book), M. Dresher, L.S. Shapley and A.W. Tucker (Eds.), Annals of Maths. Studies, Study No.52, Princeton, Princeton University Press, 1964.

67. Hoffman, A.J., and Karp, R.M., 'On Non-Terminating Stochastic Games', Management Science, Vol.12, No.5, Jan. 1966, p 359.

68. Aoki, M., Optimization of Stochastic Systems (book), New York, Academic Press, 1967.

69. Derman, C., 'On Sequential Control Processes', Annals of Maths. Statistics, Vol.33, 1964, pp 341-348.

70. Dresher, M., Games of Strategy (book), New York, Prentice Hall, 1960.

71. Berkovitz, L.D., 'A Variational Approach to Differential Games' in Advances in Game Theory (book), M. Dresher, L.S. Shapley and A.W. Tucker (Eds.), Annals of Maths. Studies, Study No.52, Princeton, Princeton University Press, 1964.

72. Radner, Roy, 'Mathematical Specification of Goals for Decision Problems' in Human Judgements and Optimality (book), M.W. Shelley II and G.L. Bryan, (Eds.), New York, John Wiley and Sons, 1964.

73. Shepard, R.N., 'On Subjectively Optimum Selection among Multiattribute Alternatives' in Human Judgements and Optimality (book), M.W. Shelley II and G.L. Bryan (Eds.), New York, John Wiley and Sons, 1964.

74. Aumann, R.J., 'Subjective Programming' in Human Judgements and Optimality (book), M.W. Shelley II and G.L. Bryan,(Eds.), New York, John Wiley and Sons, 1964.

75. Milnor, J., 'Games Against Nature' in Decision Processes (book), R.M. Thrall, C.H. Coombs and R.L. Davis (Eds.), New York, John Wiley and Sons, 1957.

APPENDIX A

SUPERCRITERIA

In this appendix we list some supercriteria to give an idea of the subjective factors that creep in. These are not exhaustive. The theory of the second best as is termed in economics is a supercriterion theory. This aspect as well as others come within the coverage of value and utility theory. A taxonomy of utility functions has been given by Radner|72|. Shepard|73| deals in detail with the notion of subjective optimality. Aumann|74| proposes alternative methods to determine subjective optimality.

In the context of our positional game Aumann's procedure is applicable in off-line decision-making when W_q is a collection of finite number of parameters or is a continuum then other methods have to be considered. In classical decision procedures under uncertainty there have been four subjective criteria very widely described in literature through they have been subject to quite some criticism. We list these for our positional game:

1. <u>Laplace's Equal Likelihood Criteria:</u>

If it is not known <u>how</u> likely a particular strategy of nature exists, it is assumed that all strategies are equally likely. Let $p_i(W_q)$ be the uniform distribution over the set W_q. Then find the function

$$E\left(I^i(\underline{u}^*) \right) = \int I^i(\underline{u}^*)dp_i\,(w) \quad u^* \epsilon \underline{U}_o \qquad (A.1)$$

and choose that $\underline{u}^*$ which minimizes this.

2. Wald's Minmax Criterion:

This minmax criterion and its wide use in engineering under worst case design practices has been one of the sources for the study of positional games.

Let $dF(w_q)$ denote a possible probability density over W_q. Then the safest optimal control strategy is the one for which $F(w_q)$ gives rise to the worst distribution and hence the maximum loss. Denote

$$\beta(u^*;w_q) = E_F \left(I(u^*;w_q) \right) \tag{A.2}$$

Then the minmax payoff is

$$\beta^*(\bar{U}^*;w_q) = E_{F^*}\left(I(\bar{U}^*;w_q) \right)$$

$$= \min_{u^* \epsilon \bar{U}_o} \max_{F} E_F \left(I(u^*;w_q) \right) \tag{A.3}$$

where E_F stands for the expectation operator with respect to the distribution F.

3. Hurwicz' Pessimistic-Optimistic Criteria:

Select a constant $0 \leq \alpha \leq 1$ which measures the decision-maker's optimism. Further let the set $\mathcal{J}^*$ of all Values of I with the use of dominant mixed strategies be bounded above and below. For each dominant mixed strategy $u^* \epsilon \bar{U}_o$ find the Value

$$\beta(u^*;w_q) = B \tag{A.4}$$

such that B^m denote the largest and B_m the smallest of

such values for different distributions P. Now choose that dominant strategy u^o such that $\alpha B_m + (1-\alpha)B^M$ is minimised. This reduces to Wald's criteria for $\alpha = 0$.

4. <u>Savage's Criterion of Regret:</u>

Let

$$r(u^*;w_q) = \beta(u^*;w_q) - \min_{u^* \in U_o} \beta(u^*;w_q) \qquad (A.8)$$

for some fixed P. Thus the function $r(u^*;w_q)$ measures the difference between the payoff which is actually obtained and the payoff which could have been obtained if the correct P was known. Now apply the Wald's criterion to $r(u^*;w_q)$.

Apart from these, certain types of approximations also constitute supercriteria.

APPENDIX B

CONDITIONS FOR APPLYING SUPERCRITERIA

The conditions under which subjective supercriteria can be applied are stated in Milnor|45|, Harsanyi|60|, Hegade|8|. These are :

1) There is definitely a non-null subset $\bar{U}$ in U which makes $I(u)$ take different values for different members in the subset. [It is the set of all dominant strategies for every initial condition on $x(t_o)$ and starting time t_o.]

2) $\bar{U}$ does not contain elements which can be ordered, i.e., it

does not depend on ordering

3) No member in the mixed extension $\bar{\bar{U}}$ can be excluded

4) If the Value $(I(\underline{u}))^k$ converges to $I_o(\underline{u})$, $(\underline{u})^k \in (\bar{U})^k$, $(\bar{U})^k$ a sequence of dominant strategies converging to $\bar{\bar{U}}_o$ and $(\underline{u})^k$ converges to u_o then $u_o \in U_o$ the limiting dominant set

5) The addition of a new member to U_o to enlarge it does not violate the dominance of the old strategies.

APPENDIX C

EXTENSIVE GAMES

Definition (Aumann[66]) :

A game from an individual player's point of view consists of

(i) A (finite or infinite) sequence $U_1, U_2, \ldots$ of standard (measurable) spaces called action spaces

(ii) A corresponding sequence $\phi_1, \phi_2, \ldots$ of standard (measurable) spaces called information spaces

(iii) A set W_{2q} called the set of strategies of the opponents (including here the player II and chance)

(iv) A sequence of functions

$$g_i : W_{2q} \times U_1 \times \ldots \times U_{i-1} \to \phi_i$$

called information functions for each $w_{2q} \in W_{2q}$ are (measurable) transformations on $U_1 \times \ldots \times U_{i-1} \to \phi_i$

(v) A standard (measurable) spaced called the payoff space

(vi) A function

$$I: W_{2q} \times U_1 \times U_2 \times \ldots \rightarrow R$$

called the payoff function. The payoff function is again a (measurable) transformation for each fixed $w_{2q} \in W_{2q}$.

APPENDIX D

GAMES OF PERFECT RECALL

A game of perfect recall intuitively calls for the player to remember not only what he did at previous moves but also what he knew at those moves. $\bar{U}_i$ U_i denote the set of all feasible strategies for player I at any instant. Then this requirement of perfect recall is equivalent to the players remembering the sequence $\left(\frac{\phi_1}{U_1}, \frac{\phi_2}{U_2}, \ldots, \frac{\phi_{i-1}}{U_{i-1}}\right)$. This is in agreement with our definitions in chapter 2. A further aid to this remembering of past moves and knowledge is afforded through the following transformations.

<u>Definition (Aumann[66])</u> :

A game is said to be of perfect recall if there are measurable transformations

$$n_j^i : \phi_i \rightarrow U_j \qquad j < i$$

$$b_j^i : \phi_i \rightarrow \phi_j \qquad j < i$$

such that

$$n_j^i : g_i(w_{2q}, \phi_1, \ldots, \phi_{i-1}) = U_j \qquad j < i$$

and $\quad t^i_j \quad g_i(w_{2q}, \phi_1, \dots, \phi_{i-1}) = \phi_j \quad j < i .$

Such a condition is naturally met by a Markov process.

APPENDIX E

A RESULT ON CONVEX GAMES

Suppose that $M(x,y)$ is continuous in both variables and is a strictly convex function of x for each y. Let $\frac{\partial M(x,y)}{\partial x}$ exist for each y in the unit interval and each x on a closed bounded interval $[0,b]$. The solution of the game is as follows:

(i) $v = \min_x \max_y M(x,y)$

(ii) Player I has a unique optimal pure strategy x^*

(iii) If $x^* = 0$ then player II has an optimal pure strategy y^*, such that $M(0,y^*) = v$ and $\frac{\partial M(0,y^*)}{\partial x} \leq 0$.

(iv) If $x^* = b$ then player II has an optimal pure strategy y^* such that $M(b,y^*) = v$ and $\frac{\partial M(b,y^*)}{\partial x} \geq 0$.

(v) If $0 \leq x^* \leq b$ then player II has an optimal mixed strategy which is of the form

$$G^*(y) = \alpha I_1(y_1) + (1-\alpha) I_2(y_2) \tag{E.1}$$

where $I_1(y_1)$, $I_2(y_2)$ are the step function distributions at $y = y_1$ and $y = y_2$ respectively where the parameters α, y_1, y_2 satisfy the condition

$$M(x^*,y_1) = M(x^*,y_2) = v$$

$$\frac{\partial M(x^*,y_1)}{\partial x} \geq 0 \geq \frac{\partial M(x^*,y_2)}{\partial x} \tag{B.2}$$

$$\alpha \frac{\partial M(x^*,y_1)}{\partial x} + (1-\alpha) \times \frac{\partial M(x^*,y_2)}{\partial x} = 0$$

This result is due to Dresher|70|.